



Évaluation ouverte par les pairs : de l'expérimentation à la modélisation

Julien Bordier

► To cite this version:

Julien Bordier. Évaluation ouverte par les pairs : de l'expérimentation à la modélisation : Récit d'une expérience d'évaluation ouverte par les pairs. 2016. hal-01283582

HAL Id: hal-01283582

<https://hal.science/hal-01283582>

Preprint submitted on 5 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Évaluation ouverte par les pairs : de l'expérimentation à la modélisation

Récit d'une expérience d'évaluation ouverte par les pairs

Julien Bordier, docteur en sociologie

Résumé : Cet article relate le déroulement de l'expérimentation d'un dispositif d'évaluation ouverte par les pairs et de commentaire ouvert, pour des propositions d'articles à la revue en sciences de l'environnement *VertigO*, publication scientifique électronique en accès libre. Si cette expérimentation ne s'est pas déroulée sur un temps assez long (4 mois) et un corpus assez étendu (10 manuscrits) pour en tirer des conclusions quantitatives fermes, elle expose néanmoins des pistes et des réflexions concrètes sur les potentialités et les limites de l'ouverture des processus d'évaluation – au sens large – pour la publication scientifique. Se basant sur l'exemplarité de l'expérience et une observation participante en tant que secrétaire de rédaction consacré à l'évaluation ouverte, l'article propose finalement la modélisation du prototype expérimenté. Ce modèle, surnommé OPRISM, pourrait être utilisé dans d'autres cadres éditoriaux pour les sciences humaines et sociales.

Mots-clés : Evaluation ouverte par les pairs, Commentaire ouvert, Edition électronique, Innovation éditoriale, Accès libre, OpenEdition, OpenAire

Pratique centrale et sujette à débat dans le monde académique, l'évaluation par les pairs y désigne à la fois l'activité de contrôle, de validation, d'attribution, de critique exercée par la communauté scientifique sur et par elle-même. Son champ d'action est vaste: il va de l'évaluation de l'attribution d'un financement à l'évaluation de la pertinence d'un recrutement. Selon l'acception commune, le fait que le travail de la communauté scientifique soit soumis à son propre contrôle est une garantie de sa qualité scientifique, cet enjeu devient central dans un contexte international de mise en concurrence des universités et du travail de recherche¹. Dans le champ de la publication scientifique, l'évaluation par les pairs consiste à faire relire une proposition de texte par des experts qui indiquent à la direction de la publication si le texte peut être publié ou non et si des corrections doivent y être apportées. La tradition veut que les auteurs ne sachent pas qui évalue leur texte – simple aveugle – ou que ni les auteurs ni les évaluateurs n'aient connaissance des noms des uns et des autres – double aveugle. Ce travail peut représenter une part importante des activités d'un chercheur alors qu'il constitue une part invisible et le plus souvent bénévole du travail effectué autour d'une publication. Si le principe de l'évaluation est largement accepté, ses modalités sont en revanche au centre de nombreuses questions². La principale critique cristallisée autour de

1 BOURE Robert, « De l'évaluation collégiale à l'évaluation à dominante gestionnaire », Communication et organisation, 38, 2010, <http://communicationorganisation.revues.org/1391> (consulté le 11/02/2016)

2 Parmi la riche littérature à ce sujet, se reporter aux repères bibliographiques en fin de texte et voir par exemple : WICHERTS Jelte M., « Peer Review Quality and Transparency of the Peer-Review Process in Open Access and Subscription Journals », in PlosOne, 29 janvier 2016, doi : 10.1371/journal.pone.0147913

l'évaluation par les pairs concerne la compétence effective des évaluateurs. En effet, il est avéré que l'évaluation n'a pas empêché la publication de nombreux résultats erronés permettant d'appuyer un financement par exemple, ou bien, ce qui est peut-être moins grave, la publication de canulars³ ; ces derniers ayant au moins le mérite de mettre en évidence les failles d'un système qui semble inamovible. De manière générale, il est évident que l'effet d'objectivité impartiale recherché par l'évaluation par les pairs relève du mythe⁴, peut-être de manière encore plus sensible dans les sciences humaines et sociales où les résultats d'une analyse dépendent des points de vue et ne visent pas nécessairement une exactitude. Là où des points de vue peuvent s'y affronter, la critique permet-elle forcément de disqualifier ?

Proposer l'ouverture de l'évaluation, tout comme permettre une interaction scientifique autour des contenus publiés, sont des alternatives pour contrer ces biais et renouveler ces pratiques. Cela repose sur un principe simple : là où les auteurs et les évaluateurs sont censés être maintenus à distance dans une évaluation classique, l'évaluation ouverte par les pairs propose de les rapprocher. Cette alternative s'appuie sur une intuition simple émanant de nombreuses voix de la communauté scientifique. Puisque l'évaluation classique ne permet pas d'empêcher les erreurs ou les canulars dans les publications, ni de contourner les effets de conflits d'intérêt dans la mesure où l'anonymat n'est souvent que de principe, il convient de la prendre pour ce qu'elle devrait être : une manière d'améliorer des textes de manière collaborative et d'orienter qualitativement des comités de rédactions. La pratique de l'évaluation ouverte par les pairs s'est particulièrement développée dans la sphère anglophone pour le domaine Science-Technique-Médicale. Elle prend des formes différentes (*crowdsourcing review*, commentaire par les pairs, évaluation pré-publication, évaluation post-publication, *synchronous review*, *peer review by endorsement*, *editor-mediated review*) mais un principe central permet de la définir. L'évaluation ouverte par les pairs implique que les rapports d'évaluation soient publiés, accessibles, signés et qu'auteurs et évaluateurs puissent échanger à leur sujet⁵.

Cette alternative ouverte à l'évaluation classique trouve aujourd'hui un terrain favorable dans

3 Comme ce fut le cas d'articles générés par SCIGen, voir VAN NORDEN Richard, « Publishers withdraw more than 120 gibberish papers », in *Nature*, 24/02/2014, doi:10.1038/nature.2014.14763 et dans la revue *Société*, voir COULMONT Baptiste, « L'Autolib' révélatrice de la sociologie postmoderne », in *Le Monde*, 9/03/2015

4 Voir par exemple le billet de blog du journaliste scientifique Michael Nielsen à ce sujet : <http://michaelnielsen.org/blog/three-myths-about-scientific-peer-review> (consulté le 08/02/2016)

5 Pour un survol général de l'évaluation ouverte par les pairs, je me permets de renvoyer à mon propre papier, « Evaluation ouverte par les pairs : polysémie et problématiques », écrit dans le cadre de cette expérience et publié sur OpenEdition LAB, <http://lab.hypotheses.org/1453> (consulté le 08/02/2016)

l'édition scientifique électronique. Les nouvelles formes d'éditorialisation⁶ rendues possibles par la publication numérique permettent aisément d'envisager cette ouverture. Là où de telles discussions ne pouvaient avoir lieu que par échange de versions papiers interposées, il est actuellement simple et possible de permettre à ces discussions de se déployer en ligne. L'utilisateur moyen d'Internet est aujourd'hui largement amené à pratiquer de tels échanges, que ce soit dans des pratiques collaboratives ou l'utilisation des réseaux sociaux. L'évaluation ouverte par les pairs dispose donc d'un espace adéquat où se développer. Au-delà de ce simple espace, elle semble pouvoir être un pilier des politiques d'accès libre à la publication scientifique. Elle s'inscrit en effet dans les perspectives définies par la Déclaration de Berlin de 2003, consacrée à l'accès ouvert, en rendant public toute une partie du travail de recherche, important et pourtant invisible, consistant à discuter et améliorer des publications. Cela semble crucial pour les sciences humaines et sociales dont les objets de recherche peuvent concerner et nourrir la réflexion du plus grand nombre.

C'est ainsi que nous proposons de décrire, mettre en récit et questionner dans cet article une expérimentation menée dans le cadre du projet européen OpenAIRE2020⁷, par OpenEdition en partenariat avec la revue en sciences de l'environnement *VertigO* et le consortium Couperin. Cette description analytique de l'expérience nous permettra de la modéliser pour la présenter comme dispositif opérationnel d'évaluation ouverte, mais aussi de commentaire ouvert⁸. L'expérience se déroulant sur un temps court et sur un corpus réduit de cinq textes soumis à évaluation ouverte par les pairs et cinq textes ouverts à commentaires, cet article n'a pas la prétention – ni la possibilité – de présenter des résultats quantitatifs et généralisables. C'est une approche qualitative qui est ici adoptée, racontant le déroulement de l'expérience pour chercher à en tirer des éléments significatifs permettant de modéliser un prototype de dispositif d'évaluation ouverte et de commentaire ouvert. Le choix de cette approche se justifie d'autre part dans la mesure où nous postulons que les enjeux liés à l'évaluation sont étroitement liés aux personnalités et subjectivités des auteurs, évaluateurs et contributeurs, la communauté scientifique étant une communauté humaine comme les autres et donc façonnée par des rencontres de personnalités et leurs interactions. Ainsi, en prenant le parti d'une approche axée sur des exemples précis, nous voulons exprimer le besoin d'analyse au cas par cas avant de se risquer à une généralisation. Au-delà des cas évoqués, l'expérience doit pouvoir montrer si la mise en place de dispositifs d'évaluation ouverte par les pairs et de commentaire ouvert est possible et comment elle peut être menée. L'analyse qui suit s'appuie donc sur

6 Pour une définition de l'éditorialisation voir notamment: VITALI ROSATI Marcelo, « What is editorialization ? » in Sens Public, Janvier 2016, <http://sens-public.org/article1059.html> (consulté le 11/02/16)

7 www.openaire.eu/openaire2020-press-release

8 Pour consulter le dispositif in-situ: <http://vertigo.hypotheses.org/evaluations-et-commentaires-ouverts> (consulté le 08/02/2016)

l'exemplarité d'une expérimentation.

Une hypothèse préalable à la mise en place du dispositif est le besoin spécifique d'accompagnement humain dans ces nouvelles formes d'évaluation et d'interactions autour de prépublications scientifiques. Nous verrons en quoi elle est vérifiée, cela ayant pu être particulièrement constaté dans la mesure où l'auteur de ce récit analytique a occupé le temps de l'expérience la place d'un secrétaire de rédaction dédié à l'évaluation ouverte par les pairs et au commentaire ouvert. On pourrait donc écrire que l'approche ici choisie relève d'une socio-anthropologie d'une innovation éditoriale, sur un corpus réduit mais dans un environnement représentatif. Cette observation participante est enrichie par les retours sur expériences qui ont pu être collectés auprès de la plupart des participants à cette expérience, qu'ils soient auteurs ou évaluateurs. Elle s'appuie également sur des discussions de vive voix et des échanges écrits informels ayant permis de mieux cerner l'objet et le contexte de l'édition scientifique, et de la place qu'y occupe l'évaluation. De plus, deux chercheurs en sciences sociales, extérieurs au dispositif, ont été interrogés à différentes étapes du déroulement de l'expérience pour recueillir leur opinion ainsi qu'un ressenti sur l'évaluation ouverte par les pairs et le commentaire ouvert tels que pratiqués dans ce protocole expérimental. L'ensemble de ces échanges a permis de nourrir la réflexion présentée ici et la modélisation qui en découle.

La retranscription et l'analyse de cette expérience sont organisées selon trois axes. Le premier décrit les deux dispositifs – évaluation ouverte par les pairs et commentaire ouvert – mis en place, le cadre éditorial dans lequel ils s'inscrivent et les possibilités techniques mobilisées. Cette première partie se limite à une description factuelle et succincte. Le second axe, plus long, relate les résultats de l'expérience. Il s'appuie sur de l'observation participante en tant que secrétaire de rédaction dédié à l'évaluation ouverte par les pairs, les retours sur expérience des protagonistes largement cités et s'inspire de la recherche périphérique menée à cette occasion. Le troisième axe pointe les limites du protocole et cherche à proposer des solutions qui pourraient permettre de l'améliorer. En ce sens, ce troisième et dernier axe émet un certain nombre de recommandations pour la mise en place d'un dispositif d'évaluation ouverte par les pairs et de commentaire ouvert. Enfin, pour conclure, après une réflexion générale sur l'ouverture de l'évaluation, une modélisation du prototype expérimenté sera exposée. Celle-ci se propose de donner les orientations principales pouvant servir à la mise en place de tels dispositifs.

1. DESCRIPTION DU PROTOCOLE EXPERIMENTAL

1.1. Cadres éditoriaux et scientifiques

Cadre opérationnel : OpenEdition

L'expérience a été menée dans le cadre technique et scientifique proposé par OpenEdition et s'appuie sur l'articulation de deux des plates-formes développées par le portail créé par le Centre pour l'édition électronique ouverte⁹ (unité mixte de services du CNRS, de l'Université Aix-Marseille, de l'EHESS et de l'Université d'Avignon). OpenEdition est une infrastructure d'édition électronique au service de la communication scientifique en sciences humaines et sociales. Elle rassemble quatre plates-formes complémentaires consacrées aux livres avec OpenEdition Books, aux revues avec Revues.org, aux carnets de recherche avec Hypothèses et aux annonces d'événements scientifiques avec Calenda. L'ensemble de ces plates-formes reçoit 3 millions de visites mensuelles provenant du monde entier. Ici, ce sont les plates-formes Revues.org et Hypothèses qui sont mobilisées.

En effet, l'expérience concerne des textes proposés à une revue hébergée par Revues.org, *VertigO – La revue électronique en sciences de l'environnement* (vertigo.revues.org) et le *Carnet de VertigO* (vertigo.hypotheses.org), carnet de recherche de celle-ci, hébergé par Hypothèses. Chaque revue hébergée par Revues.org ne dispose pas nécessairement d'un carnet de recherche. Ces derniers ont été créés pour permettre à la communauté scientifique de faire état de l'avancée de ses recherches et de communiquer à ce sujet, de manière non-académique, mais ils sont de fait très diversifiés comme on peut le constater en consultant le catalogue de la plate-forme et les carnets eux-mêmes. En février 2016, sur les 1459 carnets hébergés sur Hypothèses¹⁰, seuls 42 sont des « carnets d'accompagnement de publication » c'est-à-dire attachés à un espace éditorial; la part la plus importante étant représentée par les « carnets de programme de recherche », au nombre de 330.

OpenEdition étant spécifiquement tournée vers les sciences humaines et sociales, utiliser ses plates-formes éditoriales pour expérimenter l'évaluation et le commentaire ouverts dans ces champs disciplinaires est particulièrement pertinent. L'articulation entre une revue et un carnet de recherche hébergés sur deux plates-formes gérées par une même structure crée d'autre part un effet de cohérence éditoriale qui ne peut être que bénéfique à la mise en place et au déroulement de l'expérience.

Cadre scientifique : *VertigO*, sciences de l'environnement

⁹ <http://cleo.openedition.org>

¹⁰ <http://www.openedition.org/catalogue-notebooks>

Comme mentionné plus haut, la pratique de l'évaluation ouverte tout comme celle du commentaire ouvert, articulée à l'accès libre des publications et prépublications, s'est particulièrement développée dans le champ disciplinaire Science-Technique-Médical et notamment dans la sphère anglophone. Ce qui s'explique, du point de vue de l'histoire des sciences, par le besoin de mettre rapidement en commun et en discussions des résultats de recherche d'une part, et, par une tradition de communication scientifique spécifique d'autre part. Le cas souvent cité en exemple, bien que ce ne soit pas à proprement parler une plate-forme de publication mais une archive ouverte de prépublications ouvertes à commentaires, est *arXiv*. Mais il existe bien des revues pratiquant l'évaluation ouverte par les pairs telles que *Atmospheric Chemistry & Physics*, *Peer J* ou *F1000 Research* par exemple. Il faut remarquer que ce sont principalement des revues faisant le choix de l'accès ouvert qui mettent en place des modalités d'évaluations ouvertes. Mais cela n'est pas une nécessité : rien n'empêche à une revue à accès payant de pratiquer l'évaluation ouverte en ne laissant à ses abonnés seuls la possibilité d'y avoir accès.

Les pratiques d'évaluation et de commentaires ouverts sont peu développées dans les sciences humaines et sociales. Les raisons sont multiples et là encore la question de l'historicité de la mise en discussion des résultats de recherche est en jeu¹¹. Mener une expérience d'évaluations et de commentaires ouverts dans le cadre d'une revue en sciences de l'environnement permet de dépasser l'effet induit par les traditions disciplinaires dans la mesure où l'analyse environnementale semble nécessairement interdisciplinaire. Les sciences de l'environnement se situent en effet à la croisée des approches biologiques et humaines et trouvent leur pertinence dans la rencontre et la mise en perspectives de ces approches. Le cadre scientifique propre à la revue *VertigO* en fait donc un terrain propice à une première expérimentation de l'évaluation et du commentaire ouverts, où les différentes traditions de communication scientifiques peuvent se rencontrer sans appesantir le dispositif. Autre fait qui peut être relevé, la place internationale de la revue *VertigO* favorise sûrement un le dispositif. Revue basée à Montréal, elle est à la croisée des aires linguistiques et donc nécessairement des traditions de communications scientifiques propres à chacune d'entre elles.

La revue *VertigO* dans OpenEdition

L'articulation entre la revue *VertigO* et son carnet de recherche offre un cadre éditorial et technique qui semble pertinent pour développer un dispositif expérimental d'évaluation et de commentaire ouverts. En effet, une revue souhaitant ouvrir ses modalités d'évaluation en la rendant publique, doit

11 PONTILLE David and TORNÉY Didier, « The Blind Shall See ! The Question of Anonymity in Journal Peer Review », in *Ada : A Journal of Gender, New Media, and Technology*, No.4., 2014, doi : 10.7264/N3542KVV

d'abord disposer de l'espace adéquat où exposer ce processus. Avec son carnet de recherche, la revue dispose d'un tel espace, sans avoir à faire appel à une plate-forme externe d'évaluation comme cela peut exister.

Le carnet de recherche de *VertigO* est d'ordinaire utilisé par la rédaction de la revue à plusieurs fins. Les appels à contributions y sont publiés, certains articles publiés dans la revue y sont ouverts à commentaires (en proposant au lecteur de basculer de la revue au carnet), l'actualité de la revue y est signalée. Ce carnet de recherche est donc avant tout un outil permettant à la revue de communiquer sur ses activités et ses publications. Si il appelle à l'interactivité en proposant aux lecteurs de partager leurs points de vue sur des articles publiés dans la revue, on constate que les réactions des lecteurs y sont plutôt rares.

1.2. Description de l'expérience

L'expérience s'est déroulée sur cinq mois, le premier ayant été consacré à une recherche exploratoire, le dernier à la rédaction du présent texte, et, a été développée suivant deux branches distinctes inspirées de ce qui existe dans la multitude de formes propres à l'évaluation ouverte. L'ouverture de ces deux branches est également liée aux types de textes dont dispose la rédaction de la revue *VertigO* : de nombreuses contributions spontanées que la rédaction peine à traiter en raison de leur nombre et de leur qualité. La décision d'ouvrir deux protocoles distincts a donc été antérieure à la mise en place du dispositif, mais elle correspond bien aux possibilités offertes par l'évaluation ouverte en général, répond à un besoin de la revue en particulier et révèle enfin un attrait de celle-ci pour l'innovation éditoriale. Chacune des parties du dispositif concerne ici la prépublication, alors que les différents modes inspirés de l'évaluation ouverte peuvent s'appliquer également à la post-publication. C'est en particulier le cas de l'ouverture à commentaires. Dans cette analyse de l'expérience, on distinguera ces deux branches mais certaines des observations peuvent s'appliquer aussi bien à l'une qu'à l'autre.

Évaluation ouverte par les pairs

La première branche est un véritable protocole d'évaluation par les pairs, où cinq textes sont publiés. Son objectif est donc de faire déterminer par des experts si ces propositions d'articles peuvent être publiées dans la revue et le cas échéant quelles corrections ou améliorations doivent y être apportées. Cette évaluation est réalisée selon une grille établie par la direction de la revue,

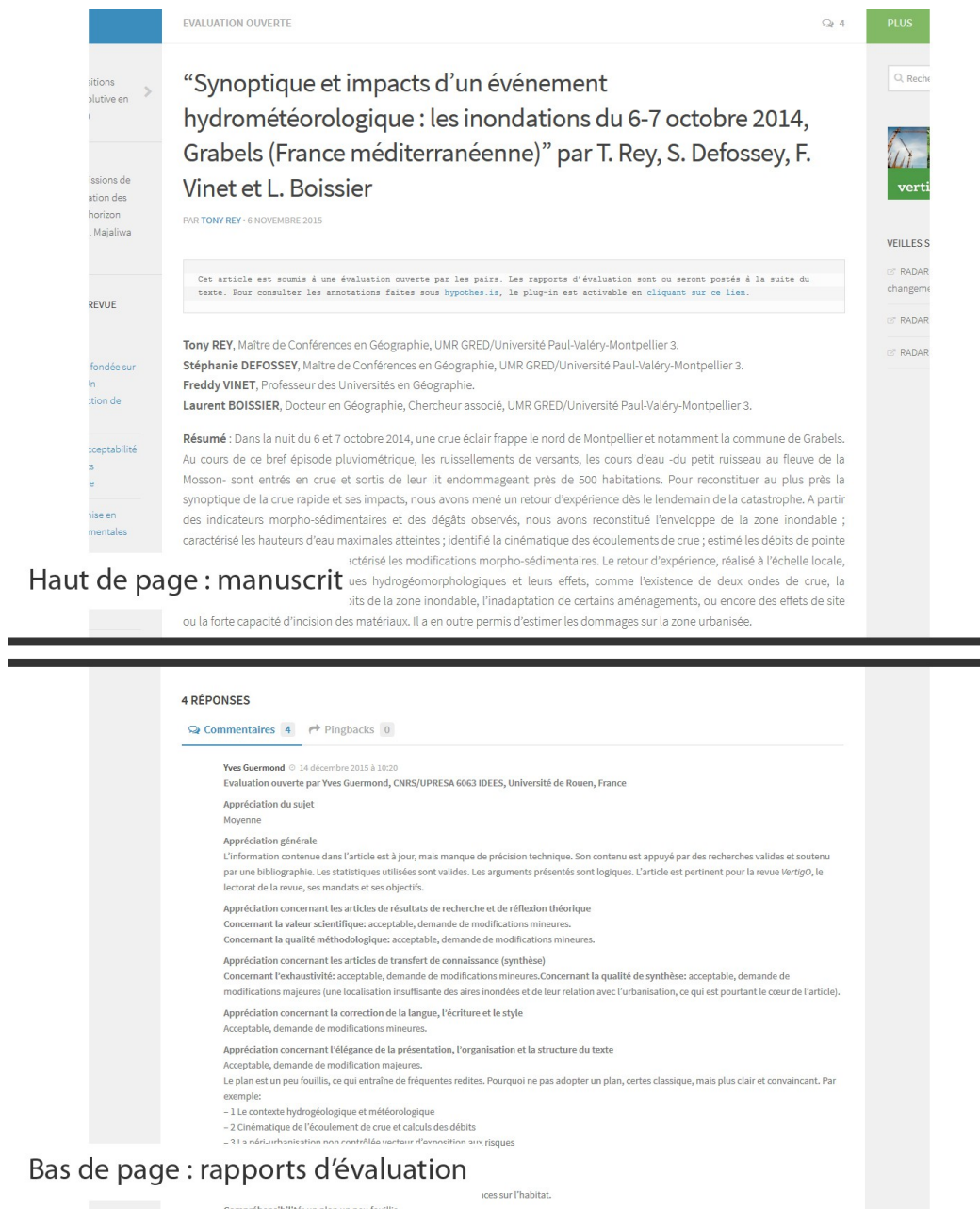
identique à celle utilisée pour l'évaluation classique¹². Les sections de cette grille couvrent les différents niveaux et caractéristiques du texte, elles sont divisées entre des parties à cocher (quatre options allant du meilleur au moins bien, en passant par deux niveaux intermédiaires, correspondant aux demandes de corrections majeures ou mineures) et des parties ouvertes à commentaires. Cette grille d'évaluation se conclut par des sections à remplir librement (points forts, lacunes, recommandations), un avis quant à la publication, le refus ou la correction majeure/mineure de l'article et enfin des observations destinées à l'auteur.

Dans le cadre de notre expérience, ce qui permet de qualifier cette évaluation d'ouverte est la mise en visibilité de l'ensemble du processus. Les textes ne sont pas anonymisés, les évaluateurs ont donc accès aux noms des auteurs. Une fois que les rapports sont remis, ils sont mis en ligne à la suite du texte et sont eux-mêmes signés. Enfin, notons que les textes sont mis en ligne dès qu'au moins un évaluateur a été trouvé pour une proposition d'article. Le contenu même de la proposition d'article est donc lui aussi rendu public dès le début du processus d'évaluation. Les évaluateurs potentiels sont désignés par la revue *VertigO*, les auteurs ont préalablement donné leur accord pour participer à l'expérience. Ici, nous nommerons cette partie du dispositif : *branche « évaluation »*¹³.

A l'issue du processus, la page d'évaluation reste en ligne dans le cas où le texte est accepté. Sur la version publiée, une note de la rédaction indique que l'article a été évalué de manière ouverte et le nom des évaluateurs. Un lien renvoie – dans une nouvelle fenêtre ou un nouvel onglet – vers la page du carnet où se situe la version originale, les rapports d'évaluation et les annotations. Sur cette page, un lien permet de faire le chemin inverse renvoyant à la version publiée, une note indiquant que le texte a été soumis à évaluation ouverte et que celle-ci est terminée. Dans le cas où le texte n'est pas accepté, le contenu rédactionnel est supprimé du carnet mais la page reste en ligne. Le titre de l'article et les noms des auteurs et évaluateurs y sont mentionnés, les rapports d'évaluation sont supprimés. Une note indique au lecteur que le texte a été évalué de manière ouverte.

12 Rendue publique parmi les « Directives aux auteurs », <https://vertigo.revues.org/5401> (consulté le 2/02/16)

13 <http://vertigo.hypotheses.org/category/evaluation-ouverte>



Bas de page : rapports d'évaluation

Illustration 1 : Un rapport d'évaluation publié à la suite d'un manuscrit

Commentaire ouvert

La seconde branche ne cherche pas à évaluer des textes en vue de leur publication. Elle se propose plutôt d'aider des auteurs à améliorer leur manuscrit avant que celui-ci ne soit soumis à évaluation, cinq textes y sont publiés, elle s'apparente à un dispositif de commentaire ouvert. Elle est ouverte à des contributions spontanées dont la rédaction de la revue estime qu'elles posent des problèmes trop importants et ne peuvent en l'état être soumises à une évaluation. Ces textes sont donc ouverts aux commentaires du plus grand nombre, il est indiqué aux contributeurs que ces commentaires peuvent

autant être de fond que de forme. L'enjeu peut donc paraître moins important dans cette partie du protocole dans la mesure où les commentaires ne vont pas décider de la destinée éditoriale d'un texte. L'expérience montre qu'il n'en est rien et démontre la pertinence de la mise en place d'un tel dispositif d'un point de vue pédagogique et du point de vue de la communication scientifique. On peut remarquer dès maintenant par ailleurs que les revues qui ouvrent leurs textes à commentaire – c'est en général le cas pour des textes déjà publiés – ne récoltent que peu de retours. Les auteurs participants à l'expérience ont donné au préalable leur accord pour la mise en ligne de leur proposition d'article. Ici, nous nommerons cette partie du dispositif : *branche « commentaire »*¹⁴ ; nous nommerons *contributeurs* ceux qui postent des commentaires et des annotations.

A l'issue du processus, les contenus rédactionnels sont dépubliés mais les pages restent en ligne indiquant le titre de l'article, les noms des auteurs et des contributeurs. Les auteurs sont invités à indiquer en remerciement les contributeurs. Dans le cas où l'article serait finalement publié dans la revue, un lien renverrait vers la page du carnet où la version initiale a été publiée, les commentaires et annotations pourraient être consultés.

The screenshot displays a web interface with a text document on the left and a comment thread on the right. The text document contains several paragraphs about coastal ecosystems and human impact. Annotations are visible in the right margin, such as "à dire environ 5,9 milliards d'hommes vit dans des zones côtières, et la plupart des grandes concentrations urbaines se situent s le littoral marin." and "Le fleuve, le littoral et les eaux marines côtières sont des milieux naturels remarquables intensivement utilisés par l'homme. Cette utilisation conduit l'ensemble des zones littorales et marines de la planète à une dégradation des ressources, une dégradation de leur environnement, une compétition pour l'utilisation de l'espace terrestre et maritime, des conflits et interférences entre des usages multiples (Prieur, 2010)." The comment thread on the right shows a user named "HerveRegnaud" posting comments and replies. The comments are dated "1 months ago".

Illustration 2 : Un contributeur annote un texte ouvert à commentaire

14 <http://vertigo.hypotheses.org/category/commentaire-ouvert>

1.3. Modalités techniques d'évaluation et de commentaire

L'expérience est mise en place sur Hypothèses, plate-forme de carnets de recherches en sciences humaines et sociales, ces derniers fonctionnant sous la forme de blogs intégrés sous WordPress. Cette forme-blog répond à une des premières exigences d'un protocole d'évaluation/commentaire puisque qu'elle permet la publication de commentaires à la suite des articles. Ainsi, déployé sur Hypothèses, un protocole d'évaluation/commentaire ouverts prend une forme simple et facilement reconnaissable par tout utilisateur : le manuscrit est assimilable à un billet de blog, les réactions qui y sont attachées sont des commentaires à ces billets.

Toutefois Hypothèses n'intègre pas la possibilité d'annoter, c'est-à-dire de publier du commentaire ligne à ligne dans un contenu. La question s'est donc posée de savoir quel serait le meilleur outil qui pourrait être proposé pour permettre aux contributeurs de soumettre des annotations dans le corps du texte. Plusieurs outils ont été envisagés comme *Framapad* et *Googledoc* notamment. Mais le premier ne permet pas d'intégrer tout ce qui est nécessaire pour une publication scientifique (comme les images et les notes de bas de page notamment) et le second pose le problème éthique de reposer sur une solution fournie par Google. Dans chacun des cas, le désavantage aurait été de faire ouvrir le texte dans une autre fenêtre ou un autre onglet à l'utilisateur qui aurait voulu saisir des annotations dans le texte. Le choix s'est donc orienté vers Hypothes.is (<https://hypothes.is>), plug-in libre et ouvert reposant sur la librairie open-source JavaScript Annotator (<http://annotatorjs.org>), qui permet de commenter ligne à ligne un contenu rédactionnel en déployant une colonne de commentaire sur une page internet (cf. illustration 2).

2. UN RECIT DES RESULTATS DE L'EXPERIENCE

On cherchera maintenant à formuler des observations et des remarques, pouvant être considérées comme les résultats du dispositif expérimental. Toutefois, comme cela a été évoqué plus haut, le corpus sur lequel s'est déroulé l'expérience n'est pas assez vaste pour en tirer des conclusions quantitatives. Mais, étayées par la recherche périphérique menée à cette occasion, ces observations sont autant de pistes pour penser les possibilités et les limites de l'évaluation et du commentaire ouverts. Dans un premier temps, des observations s'appliquant à l'ensemble du dispositif seront formulées et on s'attachera en particulier à relever les potentialités propres à l'évaluation et au commentaire ouvert. On verra ensuite quelles observations spécifiques à chacune des branches peuvent être formulées. Enfin, on s'intéressera aux problèmes apparus au cours de l'expérience ainsi

qu'aux limites – techniques en particulier – auxquelles elle a été confrontée. Cela ouvrira la voie à des recommandations qui peuvent être énoncées pour la mise en place d'un dispositif d'évaluation/commentaire ouverts de prépublications.

2.1. Résultats généraux

2.1.1. L'intérêt suscité par un renouvellement des pratiques

Échanges spontanés

Le premier point à relever est l'enthousiasme quasi-unanime vis-à-vis de l'expérience, qu'il émane d'auteurs, d'évaluateurs ou de contributeurs et des personnes sollicitées en général. Cela s'est exprimé d'abord dans la recherche d'évaluateurs/contributeurs. Comme pour une évaluation classique, un des points essentiels est de trouver des évaluateurs qui acceptent de consacrer du temps à une évaluation, celle-ci étant généralement bénévole et normalement anonyme – ce dernier point signifiant que le chercheur n'en tire pas directement profit en termes de carrière contrairement à une publication par exemple. La démarche envers les évaluateurs potentiels, désignés par la revue, s'est attachée à bien expliquer le protocole, en particulier la publication des rapports et l'identification publique des évaluateurs, dès la première prise de contact, tout en restant concis pour s'assurer que le message soit bien intégré :

« Le principe de l'évaluation ouverte par les pairs, telle que pratiquée ici, est simple. Il ne diffère d'une évaluation classique que dans la mesure où votre rapport d'évaluation et vos annotations seront publiques et signés de votre nom. »

Dans un dispositif d'évaluation ouverte, tout comme dans un dispositif d'évaluation classique, de nombreux experts déclinent la proposition¹⁵. Toutefois, les réponses négatives marquent tout de même un intérêt pour l'expérience. On peut donner quelques exemples qui semblent représentatifs :

« Je n'ai absolument pas le temps de me consacrer à ce projet qui, j'en suis sûr, sera passionnant. » ; « Le projet décrit m'a l'air intéressant mais je dois décliner la proposition faute de disponibilité de temps. » ; « Je suis dans l'obligation de décliner votre proposition, malgré l'intérêt de la recherche et le mode de correction alternative. » ; « Je trouve cette expérience très intéressante et je le ferais volontiers sauf si vous attendez une évaluation rapidement (avant la fin de l'année). » ; « L'idée est intéressante mais je dois avouer que je

15 Voir à ce sujet, cet échange entre chercheurs sur Twitter:
https://twitter.com/marc_rr/status/659318460133019648/photo/1 (consulté le 5/02/2016)

suis obligé de hiérarchiser les nombreuses sollicitations qui me sont faites. »

Où l'on peut remarquer qu'un des obstacles récurrent est le temps disponible, une remarque qui s'applique aussi bien à la recherche d'évaluateurs que de contributeurs. Ces derniers ont exprimé le même intérêt que les évaluateurs potentiels, quand bien même ils n'ont pas pris le temps de commenter des textes. A titre d'exemple on peut citer ces retours, invoquant également un manque de temps :

« Je trouve l'expérience très louable et très intéressante. J'espère trouver quelques instants pour m'y atteler mais sans garantie car je suis quelque peu "chargé" en ce moment. » ; « Merci de m'associer à cette initiative qui m'intéresse vivement, à la fois dans sa forme et dans son contenu. » ; « Cela me paraît tout à fait intéressant et essentiel, je ne suis malheureusement pas disponible en ce moment. Je tenterai de me pencher plus tard sur votre proposition. »

On ne peut dessiner le portrait sociologique type des chercheurs qui ont exprimé un intérêt, mais on peut constater qu'aucune discipline en particulier ne s'est détachée des autres. Toutefois, il me semble que ces différentes marques d'intérêt envers chacune des branches du dispositif révèle que la communauté scientifique questionne les modalités d'évaluation et n'exprime pas de frilosité particulière quant à des innovations dans ce domaine. Mais ce n'est que dans des cas exceptionnels que l'intérêt pour le dispositif a été clairement formulé comme une remise en question des modalités d'évaluation. On peut les citer dans la mesure où elles expriment bien les problèmes posés par l'évaluation classique :

« J'ai expérimenté plusieurs fois les dégâts dus à un reviewer concurrent malhonnête, ou incompetent pour des études très multidisciplinaires où les spécialistes de chaque secteur renvoient à une revue d'un autre secteur. » ; « J'ai regardé avec grand intérêt votre projet, je pense de même que l'évaluation par les pairs (ainsi que la publication) mériterait clairement d'être repensé à l'ère du numérique. »

Retours sur expérience des participants

Le même intérêt est manifesté dans les retours sur expérience qui ont été collectés auprès de la plupart des auteurs, évaluateurs et contributeurs ayant participé à l'expérience. Tous les retours sont positifs, à l'exception d'un cas dans la branche d'évaluation ouverte sur lequel nous reviendrons plus loin. Un intérêt pour de nouvelles formes d'évaluation se manifeste clairement, c'est à cette occasion qu'apparaît une critique des méthodes d'évaluations classiques. Pour des auteurs dont les textes ont été évalués par exemple :

« Pouvoir identifier les évaluateurs m'a surtout permis de juger leur bonne foi en me

permettant de me renseigner pour savoir si ils connaissaient ou non le sujet » ; « Dans une évaluation aveugle, on reçoit souvent des retours qui ne semblent pas argumentés. Ici, peut-être parce que les évaluateurs sont exposés, il m'a semblé que les remarques étaient particulièrement constructives. »

On observe le même type de positionnement du côté des évaluateurs :

« Je trouve que l'exercice est très intéressant car, [il permet] à l'auteur de mieux comprendre les remarques du correcteur en fonction du profil de celui-ci ; [il permet] de lancer le débat ; ça rend le processus "moins hypocrite" que les évaluations anonymes où souvent le relecteur comprend qui est l'auteur, mais pas l'inverse ! » ; « Sur le principe général je trouve que c'est une très bonne formule, qui permet non seulement de réaliser une évaluation, mais surtout d'avoir un retour sur celle-ci et un dialogue potentiel avec les auteurs. » ; « On se retrouve le plus souvent dans des petits cercles de spécialistes, l'anonymat est plus un principe affiché qu'une réalité. » ; « Je suis pour la levée de l'anonymat et cela a motivé ma décision de participer à l'expérience. »

Ainsi, alors que la question de l'anonymat reste un argument central pour l'objectivité des évaluations et se prémunir des conflits d'intérêts, il semble que pour les chercheurs interrogés ici l'anonymat n'est pas un facteur central à préserver dans les dispositifs d'évaluations, en ce qui concerne la publication. Si l'anonymat est en général invoqué comme moyen de protéger les parcours professionnels, il ressort de deux des citations précédentes qu'il est souvent fictif d'une part et que l'identification guiderait vers un soin particulier à la qualité des remarques d'autre part. Selon un des chercheurs non-participant, la levée de l'anonymat ne peut pénaliser professionnellement que des évaluateurs qui bâcleraient leur rapport évaluation : « *Si tu relis n'importe comment et que tout le monde le voit... tu te décrédibilises complètement. C'est aussi simple que ça.* » Le soin personnel de protéger sa crédibilité professionnelle serait garante d'un meilleur déroulement de l'évaluation. L'ouverture tendrait donc à améliorer la qualité des rapports d'évaluation. A l'inverse, on peut penser que cela peut compliquer davantage la recherche d'évaluateurs dans la mesure où ceux-ci ne sont plus protégés et peuvent finalement se retrouver jugés à leur tour. La question qui se pose finalement est de savoir si l'objectif est de trouver des évaluateurs à tout prix, sans prendre en compte la qualité de leur travail d'évaluation, ou de trouver des évaluateurs qui accepteraient de relire et évaluer des propositions d'articles sous le contrôle ouvert de la communauté.

Dans ces retours et dans la remise en cause de l'évaluation classique ressortent les avantages principaux de l'évaluation ouverte, exprimés par la communauté scientifique : la possibilité

d'échange entre auteurs et évaluateurs d'une part et, d'autre part, la levée de l'anonymat permettant aux auteurs d'apprécier les compétences des évaluateurs et donc la pertinence de leur rapport. Nous continuerons à développer ces aspects et verrons également dans quels cas ils peuvent devenir problématiques.

Statistiques et communications

L'intérêt manifesté autour de l'expérience se vérifie également dans les statistiques de visite du carnet de recherche de *VertigO*. L'annonce de l'expérience a visiblement augmenté le nombre de visites dans les jours qui l'ont suivie. La mise en place du protocole a été annoncée et mise en ligne lors de la première semaine de novembre 2015. Le premier jour de la semaine suivante, le nombre de visites était de 1215 et le nombre de pages vues de 15180 pour des moyennes respectives sur le mois courant de 395 et 1131 et sur le mois précédent de 237 et 664 (voir tableau 1, ci-dessous). De la même manière, on constate qu'à partir de novembre 2015 la page de présentation de l'expérience et deux des textes soumis à évaluations et deux des textes ouverts à commentaires font partie des dix pages et articles les plus consultés du carnet de recherche. Si ces chiffres vont par la suite décroissants et qu'ils ne nous renseignent pas sur l'idée que se font les visiteurs, cela démontre en tout cas l'intérêt – peut-être la curiosité – suscité par le renouvellement des pratiques d'évaluation. Cette hausse de fréquentation correspond également aux annonces faites sur les réseaux sociaux à propos de la mise en place de l'expérience.

	Nombre de visites	Nombre de pages vues
Premier jour de la semaine suivant l'annonce	1215	15180
Moyenne sur le mois	395	1131
Moyenne sur le mois précédent	237	664

Tableau 1 : Hausse la fréquentation du carnet suite à l'annonce de la mise en place de l'expérience

2.1.2. L'échange potentiel entre auteurs et évaluateurs

L'évaluation ouverte par les pairs permet un échange entre évaluateurs et auteurs. Toutefois celui-ci ne s'engage pas systématiquement alors qu'un travail de relance constant a été effectué invitant les auteurs à échanger avec les évaluateurs. Dans le dispositif expérimenté, on peut considérer que

l'échange prend place dans la moitié des cas, quelle que soit la branche du dispositif concernée. Il faut remarquer d'autre part que le fait que l'échange ne s'engage pas ne signifie pas que les auteurs en rejettent la possibilité. Dans le cas de la branche « évaluation », deux évaluations n'ont pas été suivies d'un échange. Cela n'est pas lié au fait que les textes soient évalués négativement : des corrections mineures sont demandées à l'un d'eux, tandis que les avis sont partagés pour l'autre (rejet d'une part et demande de corrections majeures d'autre part). Pour les auteurs, il semble plutôt que ce soit avant tout un manque de temps et une méconnaissance de cette possibilité qui les amènent à négliger cette option. Comme cela ressort d'ailleurs des échanges privés :

« Si je comprends bien, je n'ai pas à faire ces corrections, mais plutôt à en discuter avec les évaluateurs, est-ce bien cela ? C'est un processus avec lequel je suis moins habituée, mais que j'apprécie néanmoins. » ; « Je suis désolé de cette lenteur, je suis trop débordé en ce moment. [Mais] j'ai bien pris connaissance des évaluations. »

De manière générale, on remarque que si l'échange ne s'engage pas, les observations sont quand même prises en compte. Cela est particulièrement visible dans la branche « commentaires » lorsque l'on compare les versions originales et corrigées des articles et selon ce qui ressort des échanges privés entretenus avec les auteurs.

Version originale	Version corrigée après commentaires ouverts
Titre : La nappe de « Thiaroye » au Sénégal: une ressource en eau encore utilisable ?	Titre : La nappe de Thiaroye au Sénégal: une ressource en eau utilisable pour l'irrigation?
Par conséquent, cette eau est devenue inutilisable à la consommation humaine mais utilisable dans l'irrigation.	Par conséquent, cette eau est devenue impropre à la consommation humaine. Elle est actuellement utilisée dans l'irrigation non sans risques avérés.
La durée de stockage varie de un jour à sept jours. Les gobelets de consommation sont exposés à toute contamination extérieure. Ces diverses manipulations de l'eau peuvent conduire à sa contamination le long de la chaîne d'approvisionnement.	S'agissant des conditions d'usage, la durée de stockage varie de un jour à sept jours, et les gobelets de consommation sont exposés à toute contamination extérieure. In fine, ces diverses manipulations de l'eau constituent autant de sources potentielles de sa contamination le long de la chaîne d'approvisionnement.
En définitif, ces problèmes sanitaires enregistrés sont liés également à la consommation de l'eau dans le secteur	En définitif (<i>sic</i>), ces problèmes sanitaires enregistrés peuvent être associés à la consommation de l'eau dans le secteur

d'étude, montrent que les conditions de gestion de l'eau de la SONEB par la population à la base ne garantissent pas sa qualité.	d'étude, ce qui démontre que les conditions de gestion de l'eau de la SONEB par la population à la base ne garantissent pas le maintien de sa qualité.
--	--

Tableau 2 : Exemples de corrections après les retours des contributeurs

Lorsque l'échange a lieu, il prend en général la forme d'une réponse point par point aux observations qui ont été faites. Les auteurs apportent directement leurs réponses aux évaluateurs en indiquant la manière dont ils modifieront leur texte. Ces retours sont rédigés soit en réponses aux commentaires soit à l'aide de l'outil d'annotation. Dans le premier cas les auteurs tentent de reprendre dans l'ordre les remarques qui leur ont été adressées, en utilisant la fonction de réponse à un commentaire du carnet de recherche. Dans la branche « évaluation », des auteurs utilisent les entrées de la grille d'évaluation et les reprennent à leur compte pour réagir aux remarques des évaluateurs.

Exemple en réponse de commentaire : Dans le rapport : « Concernant la qualité méthodologique : Il aurait été fort à propos de présenter les techniques de cueillette du discours des acteurs dans le texte. Ce qui n'a pas été fait malgré que l'on fasse parler les acteurs. (Acceptable, demande de modifications mineures) » Dans la réponse : « Concernant la qualité méthodologique: Nous avons ajouté ces éléments méthodologiques dans le corps du texte. »

Exemple en réponse dans les annotations : Évaluateur : « La date serait utile. » Auteur : « Oui effectivement, en attendant la modification du texte, voici la précision : le livre bleu des engagements du Grenelle de la mer date de juillet 2009. »

L'échange permet également aux auteurs de moduler la pertinence d'une observation tout en y répondant. Sur une remarque concernant le manque de lisibilité sur une figure par exemple (« *une localisation insuffisante des aires inondées et de leur relation avec l'urbanisation, ce qui est pourtant le cœur de l'article* »), les auteurs répondent :

« La représentation de la crue (...) apparaît clairement dans la figure 6. Les 570 foyers touchés par les eaux apparaissent sous la zone inondable. Le lien qui apparaît graphiquement sera plus explicite dans le texte. »

Enfin, on peut noter que certains échanges se déroulent comme des discussions prolongées, les réponses amenant des nouvelles questions et précisant la remarque. Cela est un avantage dans la mesure où les auteurs ne se retrouvent pas désemparés face à une remarque qui ne serait pas assez explicite, et, cela permet à l'évaluateur de savoir si sa remarque a bien été comprise, pour la préciser

si besoin est. Ce type d'échange prolongé est davantage facilité par l'outil d'annotation que par les commentaires de type blog dans la mesure où ceux-ci sont plus longs et abordent de nombreux sujets en même temps. On peut citer cet exemple :

Évaluateur : « Que recouvre cette population locale ? », Auteur : « Nous entendons par "population locale", la population qui réside ou travaille sur un territoire telle qu'une commune ou un ensemble de communes. Mais peut être que la question visait à en savoir plus sur la composition de cette population ? », Évaluateur : « C'était plutôt la composition de cette population », Auteur : « Nous comprenons la question : sont considérés ici tant les acteurs économiques que la société civile, par opposition aux élus. Nous allons préciser les choses. »

En fonction des échanges observés ici, une typologie de base pourrait être esquissée, où apparaissent trois types principaux que sont la modification, la demande de précision et la justification.

- **Modification** : l'auteur prend acte de l'observation. Il propose la modification et/ou annonce qu'elle sera intégrée.
- **Précision** : l'auteur demande des précisions sur la correction demandée et/ou la remarque énoncée.
- **Justification** : l'auteur légitime sa position à l'appui d'arguments.

En fin de processus, le contenu de ces échanges peuvent permettre à un comité de rédaction, ou de lecture, d'arbitrer les points de vue des uns et des autres et d'orienter en fonction ses demandes de corrections. La possibilité d'échanger autour de l'évaluation ou du commentaire enrichit donc le procédé. Davantage qu'un jugement unilatéral sur le texte, l'évaluation et le commentaire ouverts sont un processus dynamique. Celui-ci n'est pas sans rappeler l'échange intellectuel qui peut avoir lieu dans un séminaire, lorsque la communauté scientifique est dans le face à face. C'est d'ailleurs ce qui ressort dans certaines prises de positions issues du retour sur expérience, notamment lorsqu'un auteur qui a engagé l'échange déclare :

« Pour ma part, j'ai trouvé la discussion intéressante. Ceci nous a permis, au-delà de l'évaluation, d'avoir des indications sur des références bibliographiques associées à des pistes à explorer, ce qui nous aidera à enrichir le contenu : ceci n'aurait pas été aussi loin dans une évaluation classique. Par ailleurs, l'évaluateur est souvent quelqu'un qui travaille dans le même domaine : ceci crée un lien, que j'aimerais dans le cas présent poursuivre. »

Enfin, on doit mentionner que si l'échange ne s'est pas engagé autour de certains textes cela est

sûrement lié à son caractère facultatif. En effet, dans le cadre de cette expérience, étant donné que les textes étaient également évalués de manière classique, certains auteurs ont sûrement préféré attendre un échange avec la rédaction autour d'une évaluation définitive que d'engager un échange autour du dispositif expérimental. On peut donc penser que si seule l'évaluation ouverte avait été pratiquée, les conditions auraient été plus favorables pour qu'un échange puisse s'engager.

2.1.3. La question de la tonalité du langage

Cette possibilité de discussion pose la question du ton sur lequel échanger. Il apparaît que l'ouverture encourage non seulement la qualité de fond mais aussi de forme¹⁶ de l'évaluation, ce qui a déjà été observé par ailleurs¹⁷. On peut relever qu'un seul échange s'est avéré problématique – nous reviendrons plus loin sur ce cas qui s'avère particulier – et qu'à trois reprises s'est posée la question pour les évaluateurs et les contributeurs de la manière dont ils devraient adresser leurs remarques. Cela a été exprimé dans des échanges privés. Dans les trois cas, un cas pour la branche « évaluation » et deux cas pour la branche « commentaire », ce sont les évaluateurs/contributeurs eux-mêmes qui ont confié une gêne quant à savoir comment exprimer leurs critiques d'un manuscrit. Ces remarques commencent par ailleurs à faire émerger la différence fondamentale entre les deux branches de l'expérience, mais aussi peut-être le fait que leur distinction n'ait pas été très claire, puisque l'on verra que cela a conditionné le ton de l'échange selon les contributeurs. Citons d'abord le cas tel qu'il s'est posé dans la branche « évaluation » :

« J'ai un vrai souci ! En effet, même si le sujet m'intéresse, je ne le trouve pas bon du tout : Mais c'est peut-être justement parce que je connais bien le sujet, que j'ai beaucoup de critiques à faire. Il y a à mon avis des graves problèmes à la fois sur le fond et sur la forme. Soit je suis franche et je dis tout ce qui ne va pas, mais c'est un peu dur pour l'auteure, soit je reste plus nuancée mais du coup je ne sers pas à grand' chose... Que faire ? »

Dans ce cas, c'est bien parce que l'évaluation est ouverte que l'évaluateur se pose la question de la formulation de ses remarques. Cette réflexion implique que dans le cas d'une évaluation classique, l'évaluateur laisse le soin à la rédaction de la revue de synthétiser les critiques adressées et que la médiation éditoriale atténuerait les critiques. Mais cette supposition est fautive puisque dans le cas de *VertigO*, comme cela semble être le cas dans la plupart des publications, les rapports

16 La question de la cordialité dans la communication scientifique est une question qui se pose par ailleurs comme le montre par exemple cette recommandation de PlosOne mis en ligne à l'adresse <http://blogs.plos.org/everyone/2015/05/01/plos-one-update-peer-review-investigation> (consulté le 2/02/2016)

17 Voir par exemple: WALSH Elizabeth, ROONEY Maeve, APPLEBY Louis, WILKINSON Greg, « Open peer review: a randomised controlled trial », in *The British Journal of Psychiatry* 176, 2000, doi: 10.1192/bjp.176.1.47

d'évaluations sont remis aux auteurs en étant simplement anonymisés. Cet épisode révèle donc un élément intéressant : l'évaluation ouverte laisse le soin à l'évaluateur d'être le propre médiateur de ses remarques. Cela permet de palier au manque de standardisation des pratiques d'évaluation. Dans une pratique ouverte où un échange est possible entre les auteurs et les évaluateurs, les auteurs peuvent davantage apprécier leurs marges de manœuvre vis-à-vis des demandes des évaluateurs.

L'évaluation ouverte personnalise donc les échanges, là où ils étaient dépersonnalisés. Cela risque-t-il d'émousser la critique ? On peut au contraire penser que cela peut la rendre plus pertinente et constructive : d'une part elle peut se déployer dans un échange, d'autre part elle se doit d'être particulièrement argumentée. On retrouve ici l'effet « séminaire » évoqué plus haut, rien n'empêche un échange critique et cordial de se déployer, d'autant plus lorsque les uns et les autres connaissent leurs champs d'expertise respectifs. Si on peut objecter que cela introduit finalement un autre effet pervers lié aux positionnements de carrière, il reste que l'ouverture – et partant la possibilité pour tous de consulter l'échange – doit pouvoir annihiler les positions de mauvaises fois ou d'abus d'autorité. En effet, de telles positions injustifiées devenant publiques sont alors visibles au grand jour et ainsi potentiellement condamnables. Dans l'exemple cité, la situation s'est réglée après un court échange avec l'évaluateur, l'invitant simplement à trouver la bonne tonalité langagière pour exprimer ses critiques. L'évaluation rendue est finalement négative mais particulièrement détaillée et argumentée. Si cet épisode ne permet pas de généraliser dans la mesure où la personnalité même de l'évaluateur reste un facteur important – ici l'évaluateur ne veut pas être « *dur* » – on pourrait extrapoler au-delà de ce cas que l'ouverture de l'évaluation pousse l'évaluateur à détailler au maximum une évaluation négative pour la légitimer¹⁸, ce qui amplifie la valeur pédagogique du rapport pour l'auteur.

Cette question s'est également posée dans la branche « commentaire ». Dans les deux cas, cela a révélé le même type de problème mais a également souligné la difficulté de distinguer les deux branches du dispositif. Cela est sûrement lié à un manque de clarté, mais aussi à une méconnaissance relative de tels dispositifs. Comme expliqué plus haut, la branche « commentaire » n'a pas pour objectif d'évaluer les textes mais d'améliorer leur qualité. Ses textes ont été choisis par la rédaction dans la mesure où ils n'auraient pas pu être soumis à évaluation. C'est bien ce que révèlent également ces contributeurs dans un échange privé. Pour le premier :

« Je suis super embêté : l'article est très faible (litote) et si j'écris les commentaires en direct ça

18 Élément déjà mentionné et également mis à jour par : WALSH Elizabeth, ROONEY Maeve, APPLEBY Louis, WILKINSON Greg, « Open peer review: a randomised controlled trial », in *The British Journal of Psychiatry* 176, 2000, doi: 10.1192/bjp.176.1.47

ne sera pas sympa du tout pour les auteurs. Ça montrera qu'ils parlent de processus qu'ils ne connaissent pas, et qu'ils citent n'importe quoi sans avoir cherché à identifier les articles qui seraient utiles pour leur terrain particulier. En ce sens une évaluation non publique est préférable. (...) J'ai plutôt tendance à accepter les papiers quitte à leur suggérer des modifs nombreuses et je ne refuse quasi jamais un papier. Mais celui-là, c'est pas possible... Je vous envoie le texte avec mes remarques, en rouge, dites-moi si vous pensez que je peux les rendre publiques ou pas. »

Où l'on comprend que le contributeur ne distingue donc pas véritablement l'évaluation du commentaire prépublication. Les remarques du contributeur sont en effet exprimées sans nuance : aucune phrase construite et une utilisation crue de la ponctuation. Là encore, c'est en échangeant avec le contributeur – autour de la tonalité du langage et de la fonction pédagogique du dispositif – que la situation s'est débloquée (« *Vous avez raison, je vais retravailler ma façon de m'exprimer pour que ce soit plus soft et moins brutal* »). Toutefois, face au manque de temps du contributeur et l'intérêt de ses observations, c'est finalement le travail de secrétariat de rédaction qui a pris en charge la reformulation et la publication des remarques, sous le nom du contributeur, avec son accord et sous son contrôle.

La seconde remarque concernant la tonalité du langage révèle le même problème, bien qu'elle montre une meilleure compréhension du dispositif de commentaire ouvert. Elle exprime toutefois la même gêne sur la manière de s'exprimer en prenant en compte la publicité des retours.

« Si on avait été dans un contexte de lecture habituel et anonyme, j'aurais dit fermement que [l'article] n'était publiable qu'avec des modifs majeures de fond, probablement non réalisables par les auteurs. »

Ainsi, dans ce second cas, c'est parce qu'il ne s'agit pas d'une évaluation à proprement dit (« *habituel* »), mais de commentaire, que le contributeur ne s'est pas exprimé « *fermement* ». On comprend d'autre part dans sa remarque que la question de l'anonymat a influencé la manière dont il s'est exprimé.

Le déroulement de l'expérience montre qu'un échange peut être cordial tout en étant critique, qu'il convient simplement de trouver les formulations adéquates, ce que l'ouverture du processus pousse à faire. On a vu également qu'un travail d'accompagnement est ici nécessaire pour débloquer des situations liées à ce sujet. L'ouverture des dispositifs d'évaluation – et la familiarisation avec de tels dispositifs – ne peut donc qu'améliorer la qualité de la communication au sein de la communauté scientifique. On peut même penser qu'elle permet de tisser du lien là où il n'y en a pas encore.

2.1.4. Utilisation de l'outil d'annotation

Comme précisé plus haut, l'expérience permet aux évaluateurs/contributeurs d'utiliser un outil d'annotation pour saisir des observations à même le texte (le logiciel Annotator développé dans une forme de plug-in pour navigateur par Hypothes.is). Une dernière constatation d'ordre général concerne la difficulté à amener certains des participants à utiliser cet outil, pourtant pertinent pour discuter d'un contenu rédactionnel en ligne. Dans la branche « évaluation » par exemple, dans chacun des cas ce n'est qu'un seul des évaluateurs qui a utilisé l'outil d'annotation. Cette faible utilisation est étonnante dans la mesure où dans la pratique habituelle d'évaluation classique, les évaluateurs renvoient en règle générale à la revue un document texte annoté¹⁹. La pratique n'est donc pas nouvelle, c'est l'utilisation de l'outil qui l'est. Mais on remarque que ceux qui n'ont pas utilisé l'outil d'annotation n'ont pas non plus donné de suite au retour sur expérience dans le cas de la branche « évaluation »²⁰. On peut donc penser que c'est la question du temps accordé à l'évaluation qui est ici en jeu, en relation avec le besoin relatif de se s'approprier un nouvel outil. Dans le cas de la branche « commentaire », les retours dont nous disposons laissent à penser que la fonction de commentaire de billet de blog était suffisante pour ceux qui s'en sont contentés. Un seul des contributeurs a utilisé les deux fonctions de façon importante.

Nourrissant l'hypothèse du temps accordé à l'évaluation et au commentaire, comprendre et installer un nouvel outil demande du temps et un effort. Ici, utiliser Hypothes.is nécessite une démarche technique de la part de l'utilisateur. Ce dernier doit créer un compte puis l'activer, activer le plug-in sur son navigateur, comprendre comment saisir une annotation, finalement saisir l'annotation pour enfin cliquer pour la publier. Pour accompagner les utilisateurs un tutoriel a été mis en ligne, dont la version vidéo a été visionnée environ 90 fois à la fin janvier 2016. Le tutoriel a été consulté par les participants en général, ce que confirme le retour sur expérience, mais cela s'avère insuffisant. La majeure partie de l'accompagnement technique a concerné l'installation et l'utilisation d'Hypothes.is, par courrier électronique et par téléphone. Il semble évident que le fait d'avoir à passer du temps sur un dispositif technique que l'on ne maîtrise pas, et pour ceux qui disposent d'une connaissance réduite des outils informatiques, un tel outil n'est pas attractif et cela constitue une véritable limite dans le dispositif. On peut donc conclure que de nombreux évaluateurs et contributeurs ont

19 Ce qui n'est toutefois pas systématique comme nous avons pu le constater pour les quelques textes dont nous avons eu accès également à l'évaluation classique.

20 Ce qui peut laisser à penser que certains estiment déjà, à raison, faire une faveur à une publication en évaluant une prépublication et qu'il est légitime de ne pas y accorder encore plus de temps en répondant ensuite à un questionnaire.

simplement abandonné l'idée d'utiliser l'outil d'annotation pour contourner la difficulté.

Mais ce constat ne doit pas pour autant faire négliger l'utilisation qui a été faite de l'outil d'annotation, puisque celui-ci a tout de même été utilisé sur tous les textes des deux branches. Les participants qui ont utilisé l'outil d'annotation semblent d'ailleurs l'avoir plutôt apprécié :

« Je n'ai eu aucun souci avec Hypothes.is [et] je ne connaissais pas cet outil » ; « L'outil est assez simple d'utilisation et se prête bien à l'exercice », « Il m'a fallu un peu de temps pour bien comprendre, mais l'outil est vraiment parfait », « Je n'ai pas compris comment l'installer, mais je me suis débrouillé avec les liens que vous m'avez envoyé. C'est en effet très pratique de pouvoir annoter l'article pour être précis », « C'est un bon moyen d'échanger avec l'évaluateur sur des points précis. »

Enfin, tous les textes de la branche « commentaire » provenant d'auteurs résidant en d'Afrique subsaharienne, il faut noter comme ils l'ont eux-mêmes fait remarquer, que les mauvaises conditions d'accès à Internet ne favorisent pas l'utilisation d'un tel outil (*« Notre difficulté a résidé dans la connexion internet, [nous n'avons pas] réussi à répondre aux annotations » ; « Avec mon internet très lent, j'ai préféré répondre par commentaires qu'en utilisant Hypothes.is. »*). En effet, le plug-in nécessite une connexion plutôt rapide puisqu'il fait transiter plus d'informations sur une même page. D'autre part, il nécessite dans un premier temps de le pratiquer pour bien en saisir le fonctionnement, ce qu'une mauvaise connexion ne favorise pas. Cela ressort clairement du retour sur expérience, comme des échanges privés qui ont eu lieu avec les auteurs de la branche « commentaire ».

De manière générale, outre cette question d'accès, quelques problèmes techniques ont été rencontrés qui ont tous été réglés en échangeant avec les utilisateurs et le cas échéant en se tournant vers le support technique d'Hypothes.is. Cela montre d'une part que l'accompagnement humain d'un tel dispositif comprend une partie technique et d'autre part l'intérêt qu'il y aurait à intégrer un tel outil directement sur l'interface où se déroule le processus d'évaluation/commentaire.

2.2. Résultats spécifiques à la branche « Évaluation ouverte »

2.2.1. Cinq situations différentes

Le principal résultat significatif concernant la branche « évaluation » est que chacun des textes a

donné lieu à des évaluations différentes. Cela laisserait entendre que l'ouverture de l'évaluation n'implique pas une plus grande acceptation ou un plus grand rejet des textes, bien qu'encore une fois on ne puisse ici établir de véritable règle quantitative étant donné le faible nombre de manuscrits soumis à l'expérience. On peut toutefois énumérer les cinq cas de figure pour en relever la diversité. Un texte a été rejeté (mais évalué par un seul pair), un texte a été accepté pour lequel les deux évaluateurs ont demandé des corrections mineures, un texte a été rejeté par un des évaluateurs et mal évalué par l'autre demandant des corrections majeures, un texte a été accepté avec des demandes de corrections majeures d'une part et mineures d'autre part, un texte a connu le même cas de figure à la différence près que l'appréciation « corrections mineures » renvoyait à de nombreuses remarques alors que l'appréciation « corrections majeures » renvoyait à peu de remarques profondes. On pourrait synthétiser ces retours dans le tableau suivant :

	Accepté sans modification	Demande de corrections mineures	Demande de corrections majeures	Refusé
Texte 1 (un seul évaluateur)				Évaluateur 1 (ou demandant des corrections majeures)
Texte 2		Évaluateur 1 Évaluateur 2		
Texte 3			Évaluateur 1	Évaluateur 2
Texte 4		Évaluateur 1	Évaluateur 2	
Texte 5		Évaluateur 1 (mais renvoyant à un important nombre de remarques et impliquant des corrections majeures)	Évaluateur 2 (mais renvoyant à peu de remarques et impliquant des corrections mineures)	

Tableau 3 : Des évaluations ouvertes différentes.

Pour un des textes (texte 5) dont nous avons pu également disposer des deux rapports d'évaluation classique, il n'y a pas de différence notable, si ce n'est que là où l'évaluation ouverte demandait des

corrections mineures d'une part et majeures d'autre part, l'évaluation classique demandait de la part des deux évaluateurs des corrections mineures. Toutefois, ce texte est celui où les demandes de corrections « majeures » et « mineures » renvoient à des appréciations qui semblent contradictoires.

Texte 5	Accepté sans modification	Demande de corrections mineures	Demande de corrections majeures	Refusé
Évaluation ouverte		Évaluateur 1 (mais renvoyant à un important nombre de remarques impliquant des corrections majeures)	Évaluateur 2 (mais renvoyant à peu de remarques impliquant des corrections mineures)	
Évaluation classique en simple aveugle		Évaluateur 1 Évaluateur 2		

Tableau 4 : Comparaison entre les évaluations ouvertes et classiques sur un texte

Le résultat significatif ressortant de cette comparaison est le fait que l'évaluation ouverte est beaucoup plus détaillée dans les commentaires alors que l'évaluation classique est beaucoup plus succincte dans son rapport d'évaluation, où quasiment seules les parties à cocher sont remplies. Toutefois, dans le cas de l'évaluation classique, les deux évaluateurs ont renvoyé un document annoté, contenant principalement des remarques de forme. Dans l'évaluation ouverte au contraire, un seul des évaluateurs a annoté le document, tandis que le second s'est concentré sur son rapport.

Texte 5	Évaluation ouverte	Évaluation classique
Qualité des rapports d'évaluation	454 mots Rapports commentés et argumentés. Large place accordée aux parties rédigées.	138 mots Seules les parties à cocher sont remplies, rares commentaires. Parties rédigées vides pour la plupart.
Annotations	12 annotations	81 annotations

	Un seul des évaluateurs utilise l’outil d’annotation.	Les deux évaluateurs renvoient un document annoté.
--	---	--

Tableau 5 : Comparaison des retours des évaluateurs sur un même texte en fonction des deux types d'évaluation.

Ces constatations peuvent être partagées pour un second texte dont nous disposons également de l'évaluation classique (texte 1). On observe que l'évaluation ouverte (mais par un seul évaluateur) propose de ne pas publier le texte – ou des corrections majeures – expliquant ce choix par un argumentaire détaillé, alors que l'évaluation classique valide sa publication, demandant des corrections mineures. Avec cette seconde comparaison on constate que l'évaluation ouverte n'implique pas nécessairement une plus grande clémence vis-à-vis des propositions d'article et confirme qu'une évaluation ouverte – et négative – est particulièrement argumentée. Cela pourrait être synthétisé dans les tableaux suivants.

Texte 1	Accepté sans modification	Demande de corrections mineures	Demande de corrections majeures	Refusé
Évaluation ouverte				Évaluateur 1 (ou demandant des corrections majeures, un seul évaluateur)
Évaluation classique en simple aveugle	Évaluateur 1 (mais demandant des précisions)	Évaluateur 2		

Tableau 6 : Comparaison entre les évaluations ouvertes et classiques sur un second texte

Texte 1	Évaluation ouverte	Évaluation classique
Qualité des rapports d'évaluation	588 mots Pour un seul rapport.	359 mots Pour les deux rapports au total.
Annotations	22 annotations	8 annotations

	Pour un seul évaluateur.	Un seul évaluateur renvoie un document annoté.
--	--------------------------	--

Tableau 7 : Comparaison des retours des évaluateurs sur un même texte (texte 1) en fonction des deux types d'évaluation.

Rappelons qu'il est difficile de tirer des conclusions fermes de notre corpus. Il reste que l'évaluation ouverte laisse apparaître des cas de figures différents en fonction des textes, ne laissant pas penser qu'elle impliquerait une plus grande acceptation ou un plus grand rejet des articles évalués de manière ouverte. Tous les cas de figures qui se sont manifestés, de l'article accepté avec demande de corrections mineures, à l'article rejeté, en passant par des évaluations contradictoires. Mais il ressort clairement que l'évaluation ouverte pousse les évaluateurs à davantage détailler leurs retours.

2.2.2. Deux cas de blocages

Blocage par l'évaluateur

Deux cas ont soulevé des problèmes qui cristallisent significativement les limites de l'évaluation ouverte. Le processus a rencontré deux types de blocages, de la part d'un auteur et de la part d'un évaluateur. On ne s'attardera pas ici à nouveau sur le second. Nous l'avons déjà évoqué plus haut : un évaluateur ne sachant comment se positionner dans la mesure où il souhaitait rendre une évaluation négative. Cette situation s'est débloquée en échangeant simplement avec l'évaluateur et en lui exprimant qu'il convenait de trouver le ton juste et qu'une critique pouvait s'exprimer de manière cordiale. Ce conseil a été suivi par l'évaluateur qui a bien pris soin d'argumenter des critiques et finalement demandé des « corrections majeures ». Comme cela a été développé plus haut en traitant la question des tonalités de langage, c'est bien cet élément qui posait ici problème, à savoir comment exprimer des critiques tout en les exprimant de manière cordiale. On peut toutefois penser que c'est la nouveauté de ce type de dispositif qui mène à cette difficulté de positionnement de la part des évaluateurs.

Blocage par l'auteur

Un second cas de blocage a soulevé un problème plus urgent, survenu cette fois du côté de l'auteur. Il faut signaler que celui-ci n'a manifestement pas bien compris les conditions de l'expérience dans la mesure où il était mécontent que son article soit accessible à tous pendant son évaluation sans

pour autant avoir l'assurance qu'il serait publié. Il a fallu dans un premier temps convaincre l'auteur de laisser son texte en ligne malgré, selon ses propres mots, sa « *démonétisation* », estimant en perdre « *la propriété intellectuelle* ». Le résultat de l'évaluation ayant été négatif, l'auteur a demandé à ce que son texte soit retiré estimant ne pas avoir « *le temps de répondre à chaque contributeur, dont les commentaires sont peu constructifs et dont la compétence par rapport au sujet traité ne me paraît pas convaincante.* » Il a donc fallu le convaincre de répondre à l'évaluation, lui assurant une fois de plus qu'il ne perdait pas la propriété intellectuelle de son article²¹. A l'évaluation certes négative mais constructive, l'auteur a répondu en rejetant toutes les remarques et en disqualifiant l'évaluateur, affirmant notamment :

« L'évaluateur a lui-même publié une contribution (...) que j'avais moi-même évalué et dans un sens positif, or un même reproche aurait pu lui être formulé. Il n'est pas demandé aux contributeurs de *VertigO* de faire de la prospective, mais de dessiner des hypothèses raisonnables. Ce que fait mon article. »

Ici, on atteint bien une limite de l'évaluation ouverte, où le débat critique n'arrive pas à se déployer. On peut toutefois penser que la réaction de l'auteur aurait été la même dans une évaluation classique, à la différence qu'elle n'aurait pas été partagée publiquement. A cela s'ajoute l'agacement de l'auteur quant au fait que son texte soit librement accessible avant même d'avoir été publié. Cette situation amène finalement un blocage dans la mesure où la vive réaction de l'auteur ferme les possibilités de discussion, comme l'a confié l'évaluateur :

« Je suis, vous vous en doutez, un peu surpris de la réaction de l'auteur et de ses réponses que j'ai ressenties parfois comme un peu agressives. Je ne sais pas si c'est qu'il n'a pas compris mes remarques ou si, tout simplement, il n'accepte ni ne comprend l'intérêt de l'évaluation par les pairs. J'ai trouvé très étonnant qu'il remette à ce point tout en question, même quand je demande simplement de reformuler des phrases pour plus de fluidité dans la lecture. Il n'a pas non plus compris que je lui demande d'aller plus loin qu'un état des lieux, non pas de faire des prévisions ou de la prospective, mais tout simplement d'approfondir son analyse. La remarque sur sa non-volonté d'expliquer sa méthodologie m'a également un peu surprise. » Ce qui lui fait ajouter : « Je ne sais pas si c'est parce que c'est un article sur lequel j'ai fait beaucoup de critiques ou si c'est lié à la personnalité de l'auteur, mais je ne sais pas si à l'avenir j'accepterai de nouveau ce type d'évaluation : sans le vouloir, je me suis attiré les foudres de quelqu'un que je ne connais pas ! »

21 Une telle remarque révèle par ailleurs une méconnaissance des pratiques propres à l'accès libre dans la publication scientifique.

Ce que montre cet épisode, dont l'issue a été la dépublication du texte à la demande de l'auteur, est le risque potentiel de voir des auteurs désertier des publications qui feraient le choix de l'évaluation ouverte. Ici, c'est bien les conditions d'ouverture combinées à la personnalité de l'auteur qui conditionnent le déroulement du processus dans la mesure où l'auteur semble redouter de voir sa réputation mise en cause ou inacceptable de voir son travail questionné. Toutefois, on peut objecter à ce risque qu'un échange normalisé amène les protagonistes à chercher à ne pas perdre la face et donc à entrer en discussion. Le blocage de la situation est donc ici, selon nous, davantage lié à la personnalité de l'auteur plus qu'au dispositif lui-même. Ce type de problème pourrait être réglé par la mise en place d'une charte de bonne conduite exclusive, proposition sur laquelle nous reviendrons plus loin.

2.2.3. Le droit de réponse des auteurs

Mais ce cas reste particulier dans le cadre de l'expérience. Pour un autre texte évalué négativement (texte 3) mais dont nous ne disposons pas encore de la réponse de l'auteur, celui-ci ne s'insurge pas de la mise en ligne d'une évaluation négative. Dans un échange privé, il écrit au contraire :

« Ce n'est en aucun cas décourageant, je l'accepte et je continue à travailler. Je verrai les commentaires et je réagirai à juste mesure car je tiens à le maintenir. »

En effet, la possibilité offerte aux auteurs de répondre aux évaluateurs s'affirme comme un avantage central de l'évaluation ouverte. Cela ressort du retour sur expérience mais aussi de l'observation du dispositif. De fait, le point de vue de l'évaluateur n'est plus unique et central dans la mesure où l'auteur peut y répondre et défendre son point de vue face à une remarque qui lui paraît illégitime, comme déjà mentionné dans la typologie établie plus haut. Si les auteurs répondent en général pour indiquer en quoi ils apporteront les corrections demandées, ils peuvent également contredire les observations faites par les évaluateurs. Par exemple :

Évaluateur : « Il faudrait cartographier : les maisons, les équipements, les voies de communications affectées (...). » ; Auteur : « Bâtiments et infrastructures sont représentés dans les figures 1 et 5. »

Ou encore :

Évaluateur : « Le plan est un peu fouillis, ce qui entraîne de fréquentes redites. Pourquoi ne pas adopter un plan, certes classique, mais plus clair et convaincant. Par exemple (...). »
Auteur : « Le plan choisi par les auteurs répond à des problématiques clairement exposées en

introduction. Il s'agit ici d'un retour d'expérience qui offre l'opportunité d'acquérir de l'information à haute résolution spatio-temporelle. L'information n'est pas segmentée et disparate comme on peut le retrouver dans les rapports techniques. »

De plus, les réponses des auteurs peuvent permettre à la rédaction de la revue, mais aussi aux lecteurs en général, d'apprécier la pertinence des critiques des évaluateurs et la manière dont les auteurs sont capables de s'adapter à la critique. A cela s'ajoute que les auteurs peuvent se renseigner sur les évaluateurs et ainsi mieux comprendre la pertinence de leurs observations. C'est un élément qui ressort clairement du retour sur expérience lorsque il est demandé aux auteurs dans quelle mesure ils ont cherché à savoir qui étaient les évaluateurs une fois leurs rapports et annotations publiés en ligne :

« [Je ne me suis renseigné sur les évaluateurs qu'] aujourd'hui uniquement car compte tenu des réponses, je me doutais bien qu'il ne s'agissait pas de spécialistes » ; « Je trouve que X, par son profil, est bien indiqué pour mon sujet plus, que Y que je sais plus à l'aise dans [d'autres] domaines. »

De manière générale, comme l'a remarqué un évaluateur, « l'évaluateur devient évalué ». En effet, l'hégémonie de l'évaluateur médiatisée à la discrétion de la rédaction de la revue est remise en question par l'ouverture. Le processus d'évaluation ouverte introduit donc une dimension de réciprocité. Malgré les risques propres aux personnalités de chacun, dont l'expérience montre la réalité concrète, l'ouverture permet donc une meilleure répartition et prises en compte de la légitimité des protagonistes ; cette répartition s'opère ici à la vue de tous et donc sous le contrôle de la communauté.

2.2.4. Mesurer le sérieux du processus d'évaluation

Dans cette même optique, l'ouverture permet à chacun de constater l'effort fourni par les uns et les autres. Elle permet aux auteurs et aux lecteurs de constater le temps qu'a consacré un évaluateur à son texte, en fonction de la manière dans son rapport est complet et le texte annoté. En retour, un évaluateur peut constater de quelle manière un auteur a pris en compte ou non ses observations. La rédaction qui doit arbitrer en fin de processus si le texte peut être publié, peut également en jugeant de l'utilisation des outils d'évaluation et de la qualité de l'échange, déterminer si l'évaluation est pertinente et si l'auteur la prend bien en compte – ce qui implique également qu'une partie de la politique éditoriale devient publique. Disposer d'une telle donnée peut aider les rédactions à choisir

ou non un texte et les lecteurs à comprendre les enjeux qu'ils soulèvent.

2.3. Résultats spécifiques à la branche « Commentaire ouvert »

2.3.1. Spécificité de la branche dans le protocole

Les cinq textes ouverts à commentaires ont été choisis par la rédaction de la revue parce qu'ils présentaient avant tout des problèmes de forme. L'objectif n'était donc pas de les évaluer mais de proposer aux auteurs un dispositif leur permettant d'améliorer la qualité de leur proposition d'article. Comme on l'a vu plus haut, si la plupart des contributeurs ont bien compris le dispositif, cela n'a pas été si clair pour certains d'entre eux. On peut donc se demander si il n'aurait pas été plus pertinent de mieux distinguer les deux parties du dispositif. Tel que cela a été mis en place, les branches « évaluation » et « commentaire » sont des catégories contiguës sur le carnet de recherche, ce qui prête sûrement à confusion. On pourrait donc envisager de mieux les distinguer. D'autre part, on peut penser là encore que la nouveauté de tels procédés crée une certaine confusion.

C'est également l'origine exclusive des propositions d'articles qui marque la spécificité de la branche « commentaire », où tous les textes sont issus d'Afrique subsaharienne. Pour des chercheurs qui ont peu d'occasions d'échanger à la communauté scientifique internationale et qui manquent de ressources autant financières que documentaires, la branche « commentaire » prouve tout son intérêt. Elle ouvre en effet un espace où des chercheurs peuvent discuter et faire évoluer leur travail. Dans le cas des textes publiés dans cette section, le non-spécialiste comprend d'ailleurs aisément l'enjeu fort de leurs objets de recherche, comme les questions de qualité de l'eau ou de réduction des pollutions par exemple. Les difficultés majeures liées à cette origine des textes sont les conditions d'accès à Internet pour les auteurs. On peut en outre penser, que les auteurs eux-mêmes n'aient pas communiqué autour de cette ouverture à commentaire pour attirer des contributeurs, dans la mesure où leurs confrères locaux partagent la même difficulté technique.

2.3.2. Trouver des contributeurs

Dans le cadre de la branche « commentaire » il a été difficile de trouver des contributeurs. Cela est une donnée constante pour tous les dispositifs de commentaires ouverts, comme on peut l'observer dans les publications en accès libre ouvrant leurs contenus publiés à commentaires. Sur le carnet de recherche de *VertigO* lui-même, où une section est consacrée à la discussion d'articles publiés dans

la revue, l'utilisation de la fonction de commentaire n'est pas courante : sur 29 articles ouverts à la discussion, neuf d'entre eux sont commentés, jamais plus de deux fois²². C'est ce qu'on peut observer de manière générale sur les carnets de la plate-forme Hypothèses, où la fonction de commentaire est finalement peu utilisée, malgré son intérêt.

Dans le cas d'un dispositif de commentaire ouvert, lui-même lié à un dispositif d'évaluation ouverte, la question de la légitimité des contributeurs s'est largement posée. Quand bien même les textes ouverts à commentaires avaient besoin autant de corrections de forme que de réflexions de fond, peu de chercheurs se sont risqués à poster des commentaires. Pour tenter d'amener des contributeurs à commenter les textes plusieurs stratégies ont été mises en œuvres. La revue *VertigO* d'une part a communiqué autour de l'expérience, invitant ses lecteurs à commenter les textes, sans que cela n'apporte de résultat. A cette stratégie de recrutement de contributeurs, on peut penser là encore que la présentation conjointe des deux branches de l'expérience a conduit à une certaine confusion. En effet, l'évaluation n'est pas ouverte à tous mais à deux experts désignés, contrairement à la contribution par commentaires. Confusion qui s'est d'ailleurs manifestée en pratique dans la mesure où un contributeur est intervenu sur un texte soumis à évaluation en pensant intervenir sur un texte ouvert à commentaire.

Communiquer de manière non-ciblée

Ce sont ensuite deux campagnes de communications qui ont été lancées successivement, la première n'apportant toujours pas de résultat. Celle-ci consistait à contacter par courrier électronique des chercheurs repérés sur des annuaires de laboratoires de recherche. Mais le courrier qui leur était adressé n'était pas spécifique, il ne faisait qu'annoncer la mise en place du dispositif et inviter à y collaborer, sans cibler spécifiquement le correspondant en fonction d'un texte en particulier. Bien que cette première stratégie n'ait pas amené de contributeurs, elle a donné lieu à de nombreux retours de la part des chercheurs contactés, avec plus d'une réponse pour 10 courriels sur environ 300 personnes contactées. Ces réponses relèvent systématiquement deux points qui semblent significatifs. Il y a d'une part, là encore, l'intérêt suscité par l'expérience – « *passionnant* », « *tout à fait intéressant* », « *à développer* », « *très pertinent* », « *vraiment intéressant sur le principe* » pour citer quelques retours – et le taux de réponse plutôt élevé pour une opération de communication impersonnelle de ce type, tout en soulevant son intérêt. D'autre part, outre le manque de temps, la raison invoquée pour décliner l'invitation était le manque de connaissances spécifiques liées aux objets traités – « *je ne suis pas du tout expert en ces*

22 A la fin janvier 2016.

domaines », « *je ne me sens pas très proche des thèmes proposés* », « *je ne suis pas sûr d'avoir les compétences nécessaires* » pour citer quelques exemples. Ce second point est tout à fait intéressant dans la mesure où les chercheurs contactés auraient pu, bien que travaillant sur des sujets éloignés, proposer les corrections de forme dont ces textes avaient besoin, mettant en évidence la difficulté de séparer le fond et la forme. On peut donc difficilement ouvrir des textes à commentaires en proposant d'intervenir uniquement sur la forme. Cette première stratégie de communication met en évidence la question de la légitimité à intervenir sur un texte et la difficulté de l'échange interdisciplinaire. On peut postuler que, dans le cas où ce type d'expérience se généraliserait, la question ne se poserait pas de manière aussi systématique.

Communiquer de manière ciblée

Face à l'échec de cette première stratégie pour attirer des contributeurs, des chercheurs ont par la suite été contactés en fonction de leurs spécialités. Ainsi, au lieu d'annoncer simplement l'ouverture du dispositif, le courrier d'invitation présentait l'ouverture à commentaires d'un texte en particulier, en fonction de l'expertise de la personne contactée. La recherche de contributeurs s'apparente alors à la recherche d'évaluateurs, bien qu'il ne soit pas demandé le même travail dans les deux cas. Cette manière de procéder a porté ses fruits puisque l'ensemble des contributeurs ayant intervenu dans la branche « commentaire » ont été trouvés de cette manière, mis à part deux d'entre eux. Cela semble donc être la bonne façon de procéder pour générer de l'interaction autour de textes ouverts à commentaires. A la suite de la mise en place de cette stratégie, le nombre de contributeurs est passé de deux à dix, à des degrés d'implications différents. Cela implique une constatation importante déjà mentionnée plus haut à un autre propos, c'est le besoin d'une médiation humaine spécifique, un secrétariat de rédaction adapté, pour le succès d'un tel dispositif.

Cette stratégie met en évidence un autre élément. En effet, environ vingt personnes n'appartenant pas à la communauté scientifique ont également été contactées de la même manière, en fonction de leur connaissance professionnelle de l'objet étudié – notamment sur les textes liés à l'utilisation de l'eau. Aucun d'entre eux n'a donné de réponse. Au-delà du temps disponible, on peut se demander si c'est une question de légitimité qui se pose ou si cela est révélateur du travail profond qui doit être réalisé pour permettre un échange entre la communauté scientifique et des experts non-académiques.

On peut estimer que le dispositif a rencontré un certain succès dans la mesure où tous les textes ont été commentés, bien qu'inégalement en termes de nombre de commentaires et d'annotations. Cela

distingue le dispositif mis en place sur le carnet de recherche de *VertigO* des autres dispositifs du même type, qui peinent en général à recueillir des commentaires. On doit donc ici insister sur la nécessité d'aller chercher nommément les contributeurs si l'objectif est de recueillir de nombreux retours sur les textes ouverts à commentaires.

2.2.3. Modalités des échanges entre contributeurs et auteurs

Les commentaires postés sur les textes de cette branche ont autant été de forme que de fond, comme le remarquent les auteurs eux-mêmes dans le retour sur expérience. Si les deux possibilités d'intervention sur le texte – commentaire de type blog et outil d'annotation – ont été utilisées, c'est principalement la fonction de commentaire qui a été favorisée. Cela s'explique à nouveau par les efforts nécessaires à l'utilisation de l'outil d'annotation déjà mentionnés plus haut. On peut toutefois établir la remarque suivante : la fonction de commentaire est plutôt utilisée pour relever des remarques de fond, quand l'outil d'annotation est plutôt utilisé pour les remarques de forme (typographie, syntaxe et orthographe), bien que cette remarque ne s'applique pas à la branche « évaluation ». On remarque sur un des textes par exemple, que deux des contributeurs ont repris le plan du texte point par point pour y apporter des remarques en commentaire. D'un point de vue pratique, cela aurait été pourtant plus simple et lisible en utilisant l'outil d'annotation. Cela nous pousse encore une fois à penser qu'intégrer un outil d'annotation à même la plate-forme favoriserait son utilisation.

Tous les auteurs de la branche « commentaire » dont nous disposons d'un retour sur expérience – trois sur cinq – ont apprécié de participer à un tel échange. Tous se disent prêts à se prêter à nouveau à ce type d'exercice. Sur les cinq textes, quatre versions corrigées ont été renvoyées. Encore une fois, la seule limite évoquée concerne les conditions d'accès à internet. Une limite qui s'explique par la situation géographique des auteurs.

3. RECOMMANDATIONS ET SOLUTIONS AUX LIMITES TECHNIQUES RECONTREES

Au terme du récit du déroulement de cette expérience et de ce qu'elle a mis en évidence, trois éléments essentiels permettent de formuler des préconisations pour la mise en place de dispositifs d'évaluation ouverte par les pairs et de commentaire ouvert. Le premier concerne les moyens techniques qui peuvent être mobilisés. On peut à partir du prototype développé au cours de cette

expérience relever ses limites techniques et proposer des solutions pour les dépasser. Le second, comme on l'a constaté à différents niveaux, est le besoin d'un accompagnement humain du dispositif. On détaillera cet aspect et cherchera à estimer ce que cela représente en termes de temps de travail. Le dernier, qui n'a pas encore été évoqué mais qui s'est dessiné au long de la mise en place de l'expérience – et s'inscrit dans l'articulation entre accès ouvert et évaluation ouverte – concerne la rétribution des évaluateurs/contributeurs.

3.1. La plate-forme Hypothèses : un espace adéquat

3.1.1. Une plate-forme de carnets de recherche académique reconnue

Pertinence de l'utilisation d'Hypothèses

L'expérience a montré la pertinence de l'utilisation d'un carnet de recherche hébergé par la plate-forme Hypothèses pour la mise en place d'un tel dispositif. C'est d'abord la forme de *blog* qui en fait un espace adéquat pour deux raisons principales : sa simplicité d'administration et ses possibilités intrinsèques de commentaire des contenus²³. Il est en effet nécessaire qu'il soit techniquement simple de mettre les textes à évaluer et/ou à commenter en ligne par un secrétariat de rédaction. Cette tâche ne doit pas nécessiter de grandes connaissances techniques et ne pas mobiliser trop de temps. C'est le cas avec Hypothèses qui, lorsque le principe de fonctionnement relativement simple d'un blog WordPress est acquis, permet de mettre rapidement des textes en ligne. D'autre part, c'est bien-sûr la fonction intégrée de commentaire de type blog qui rend pertinent l'utilisation d'un carnet de recherche académique tels que ceux hébergés par Hypothèses. Enfin, bien que cela ne ressorte pas des résultats de l'expérience – puisque cela n'a pas été directement enquêté – il semble que le caractère reconnu de la plate-forme offre une certaine légitimité à ce dispositif innovant. L'articulation entre la revue et son carnet de recherche, et le fait qu'ils puissent être identifiés dans un même environnement éditorial reconnu, fournit un appui important au développement et à la légitimité du dispositif.

Cohérence éditoriale

Dans le cas particulier de la revue se prêtant à l'expérience, le fait que le dispositif d'évaluation soit hébergé sur une plate-forme du même organisme éditorial offre une cohérence, c'est-à-dire pour la communauté scientifique un environnement reconnu et rassurant. On peut d'autre part arguer que

23 Sur l'utilisation et l'appropriation scientifique des blogs, voir: DACOS Marin et MOUNIER Pierre. « Les carnets de recherche en ligne, espace d'une conversation scientifique décentrée ». in *Lieux de savoir, T.2, Gestes et supports du travail savant*, 2010, Albin Michel, Paris, pp.333-352 (disponible sur HAL: <sic_00439849>)

cela offre une pérennité à l'ensemble des contenus mis en ligne. Cela laisse à penser que le passage d'un espace (publication définitive) à l'autre (processus d'évaluation) doit être favorisé dans un souci de cohérence éditoriale – ce qui n'est pas le cas lorsque les revues font appel à un service extérieur d'évaluation. De plus, cela permet à la revue de garder la main sur son processus d'évaluation tout en disposant de l'espace adéquat – distinct mais explicitement articulé – où le mener.

Si l'expérience ne s'est pas déroulée sur un temps assez long pour aller de l'évaluation ouverte à la publication définitive et que cela n'a donc pas pu être testé, il convient en fin de processus de créer une connexion évidente entre la version finale de l'article et sa version prépubliée sur le carnet relatant le processus d'évaluation et/ou la première version ouverte à commentaires. Afficher une cohérence éditoriale entre les différentes versions semble important pour attester de la validation du dispositif par et dans un cadre institutionnel, ce qui correspond au besoin d'adaptation de la communauté à ces pratiques. Cela peut également faciliter un référencement cohérent sur des plates-formes affiliées d'une part, et faciliter un moissonnage des métadonnées dans le cas où celles-ci renseigneraient des informations concernant l'évaluation d'autre part – recommandation qui sera formulée plus loin.

La cohérence éditoriale existant dans le cadre de cette expérience entre la revue *VertigO* hébergée sur Revues.org et son carnet de recherche hébergé sur Hypothèses pourrait être reproduite dans d'autres dispositifs du même type. Sur le modèle utilisé par OpenEdition par exemple, on pourrait imaginer un menu transversal aux différentes plates-formes permettant de passer de l'une à l'autre et faciliter ainsi la consultation du processus d'évaluation.

3.1.2. Des améliorations techniques possibles

Meilleure mise en forme éditoriale

Si la plate-forme Hypothèses s'affirme comme un espace adéquat, on pourrait imaginer certaines améliorations la rendant encore plus opérationnelle. Comme l'a fait remarquer un des chercheurs non-participant interrogé sur les avantages et limites du dispositif, il serait souhaitable de pouvoir extraire aisément le texte à évaluer et à commenter. Sur le même modèle, la version annotée et commentée pourrait elle aussi être exportable. Pour la branche « évaluation », des versions des textes sont envoyées aux évaluateurs dans un souci de confort de lecture. Cette possibilité pourrait être automatisée, indiquant sur l'objet exporté que la version n'est pas définitive pour éviter toute confusion. Une autre amélioration éditoriale serait de numéroter les paragraphes des textes, comme

cela est le cas par exemple dans la structuration éditoriale proposée par Revues.org, pour que les évaluateurs/contributeurs puissent aisément renvoyer leurs observations à des parties du texte. Ces quelques améliorations techniques, dans le cas de l'utilisation du système fourni par WordPress, devraient être réalisables en travaillant en fonction le thème utilisé. On pourrait donc penser le développement d'un thème spécifique à l'évaluation et au commentaire ouverts ou l'utilisation, sur des carnets dédiés à ces tâches, du plug-in CommentPress Core²⁴ développé pour WordPress. Mais celui-ci ne facilite pas la consultation d'un commentaire long (qui reste encadré dans l'espace limité aux annotations) tel qu'un rapport d'évaluation, bien qu'on puisse en effet publier un commentaire général. Le commentaire de type blog rend au contraire cela possible. Il reste donc une formule intermédiaire à développer.

Intégration de l'outil d'annotation

Au-delà de ces quelques améliorations, et comme le permet également CommentPress Core, un développement technique primordial serait l'intégration de l'outil d'annotation à même la plateforme sur laquelle est mise en place le dispositif. L'expérience a montré la difficulté d'amener les utilisateurs à annoter les textes alors que cela s'avère tout à fait pertinent pour commenter un texte dans le détail. C'est manifestement une question de temps et le manque d'envie de découvrir un nouvel outil informatique qui a freiné les évaluateurs et les contributeurs. Outre ceux qui ont utilisé l'outil d'annotation et qui selon le retour sur expérience n'ont pas eu de problème spécifique pour l'utiliser, on peut citer ceux qui ont au contraire choisi de ne pas l'utiliser ou ont rencontré des problèmes :

« Ça a été laborieux pour moi de comprendre comment activer Hypothes.is, sans votre suivi, j'aurais simplement abandonné et me serais concentré sur le commentaire général. » (évaluateur) ; « J'y ai jeté un œil, mais ça m'a semblé trop complexe. » (contributeur) ; « J'ai consulté les annotations en suivant le lien que vous m'avez envoyé mais je n'ai pas créé de compte pour utiliser le programme. » (auteur) ; « Notre principale difficulté a été de créer notre compte et de consulter les annotations. » (auteur).

L'alternative serait donc de faciliter l'utilisation d'un tel outil en l'intégrant à la plateforme où se déroule l'évaluation et/ou le commentaire. Cette solution standardiserait également son utilisation sans être dépendant des différents navigateurs utilisés, ce qui simplifierait l'accompagnement technique en cas de problème. Elle éviterait enfin de laisser le soin à l'utilisateur de comprendre

24 Voir par exemple futureofthebook.org/commentpress ou son utilisation pour de l'évaluation ouverte par les pairs sur <http://adareview.fembotcollective.org> (consultés le 2/02/16)

comment l'activer, d'y enregistrer un compte d'utilisateur, d'activer son compte en attendant la réception d'un courriel notamment. Un outil intégré permettrait en outre aux auteurs/évaluateurs/contributeurs de disposer d'un identifiant unique sur une plate-forme, facilitant le référencement de leurs travaux d'évaluation, de leurs annotations, au même titre que leurs publications. Bien choisir l'outil d'annotation et en simplifier l'utilisation semble être une amélioration technique centrale à apporter.

Autres améliorations possibles

Deux autres améliorations peuvent être envisagées. La première, relativement simple, est de mieux contextualiser les textes en évaluation ou ouverts à commentaire. Dans le cadre de l'expérience, cette partie explicative était mise en ligne sur une page à part et une simple phrase introduisait l'expérience en chapeau de chacun des textes. Un texte explicatif plus conséquent devrait être publié en en-tête des textes. On ne peut en effet attendre des utilisateurs qu'ils aillent consulter d'eux-mêmes la documentation relative si elle n'est pas immédiatement accessible. Cela éviterait d'autre part certaines confusions sur la différence entre l'évaluation ouverte et le commentaire ouvert et continuerait de familiariser la communauté scientifique avec ce type de protocoles alors qu'ils restent encore assez rares.

Dans le cadre de l'expérience mise en place, c'est le secrétaire de rédaction qui met en ligne les rapports d'évaluation remplis par les évaluateurs²⁵. On pourrait envisager que cette tâche soit automatisée. L'évaluateur pourrait remplir la grille d'évaluation en ligne pour qu'elle soit ensuite automatiquement traitée et mise en ligne à la suite de l'article évalué. Dans un tel cas, on pourrait distinguer les parties du rapport qui seraient publiées en ligne et celles qui ne resteraient accessibles qu'à la rédaction de la revue – informations personnelles et observations à l'intention de la rédaction seulement. Il est à remarquer que si il n'était prévu que de publier certaines parties de la grille d'évaluation au début de l'expérience, il s'est avéré nécessaire de publier l'ensemble de la grille et de retranscrire les parties à cocher. En effet, ne publier que certaines des appréciations du rapport d'évaluation ne fournissaient finalement pas assez d'informations pertinentes aux auteurs.

Recommandations 1:

- **Trouver les environnements éditoriaux adéquats – de la revue à son dispositif d'évaluation ouverte – offrant une légitimité et un appui au dispositif, et disposant**

25 Pour information, on peut consulter la grille d'évaluation telle qu'elle est établie par la revue *VertigO* : <https://vertigo.revues.org/5427?file=1>

d'une forme aisément identifiable par la communauté.

- **Permettre l'identification unique des utilisateurs, renvoyant autant à leurs publications qu'évaluations et commentaires.**
- **Intégrer l'outil d'annotation directement sur l'espace où se déroule l'évaluation et/ou le commentaire ouverts.**
- **Favoriser l'accessibilité, le confort de lecture et faciliter la citation du texte – exportation du contenu, numérotation des paragraphes par exemple.**

3.2. Un dispositif qui doit être accompagné au-delà de l'outil technique

L'expérience a clairement confirmé l'hypothèse formulée au départ : la seule présence de dispositions techniques ne suffit pas à rendre des dispositifs d'évaluation ouverte par les pairs et de commentaires ouverts efficaces et véritablement interactifs. C'est d'abord un travail important d'explication du dispositif qui est essentiel et requis pour l'ensemble des protagonistes. Comme l'a montré l'expérience, la plupart des évaluateurs ont demandé à en savoir plus sur le dispositif avant d'accepter d'y participer. Cela s'explique avant tout par la nouveauté de ce type de dispositif. On a vu d'autre part dans le cas du blocage de la part d'un auteur que cette explication n'avait pas été suffisante ou assez explicite : l'auteur ne comprenait pas pourquoi son texte était en ligne et en a demandé la dépublication.

3.2.1. Guider la bonne tenue des échanges et l'efficacité du dispositif

On a également signalé qu'il a été nécessaire, dans certains cas, de guider les protagonistes des deux branches pour qu'ils trouvent les tonalités de langage adaptées. Un secrétariat de rédaction disponible et disposant de quelques notions de médiation est donc nécessaire. Sur cette question, comme cela existe dans certaines revues, une charte de bonne conduite dans les échanges entre auteurs et évaluateurs/contributeurs, pourrait être rédigée. Celle-ci pourrait être une base à fournir aux protagonistes et pourrait permettre à un comité de rédaction de s'y référer et éventuellement de « bannir » ceux qui ne la respecteraient pas. Un tel fonctionnement semblerait plus juste qu'un système de points qui prendrait le risque de devenir trop instrumental. Avec une charte de bonne conduite et un référencement efficace des processus d'évaluation et/ou de commentaire, il suffirait de consulter un historique des références – et une éventuelle notification à un non-respect d'une charte – pour décider ou non de faire appel à un évaluateur/contributeur.

Enfin, l'utilisation des différentes fonctionnalités doit être facilitée par une activité de support technique. Si la fonction de commentaire n'a pas posé de problèmes, une assistance technique a été nécessaire pour de nombreux protagonistes quant à l'utilisation de l'outil d'annotation. On ne peut donc postuler que l'évaluation ouverte par les pairs et l'ouverture à commentaires puissent faire l'économie d'un travail de suivi du dispositif. Elles appellent au contraire des activités de secrétariat de rédaction spécifiques, en particulier tant que la communauté scientifique n'est pas familière de ce type de dispositifs. Une facilitation et une médiation humaine doit permettre de gérer efficacement cet effet de nouveauté.

Facilitation pour la branche « évaluation »

Dans le cas de l'évaluation ouverte, un travail de relance des évaluateurs et des auteurs est nécessaire. Il ne s'agit pas simplement de relancer les évaluateurs, comme pour une évaluation classique, il convient également d'avertir les auteurs de la mise en ligne des rapports et des annotations et de les inviter à y réagir. Si l'échange peut commencer à se déployer de lui-même une fois engagé, il convient pour le secrétaire de rédaction d'inviter à l'amorcer. Dans le cadre de l'expérience, il a été nécessaire de signaler aux auteurs qu'ils étaient invités à réagir à l'évaluation pour que l'échange soit effectif.

Facilitation pour la branche « commentaire »

Dans le cas de la branche « commentaire », on a bien constaté que la mise en ligne de textes et l'annonce de leur ouverture à commentaires ne suffisaient pas à attirer des contributeurs. La tâche la plus importante à assurer est alors la recherche de contributeurs. Il convient de démarcher des auteurs, chercheurs et/ou professionnels concernés par l'objet du texte publié, trouver comment les contacter et les inviter à commenter le texte en explicitant le dispositif. De ce qui a pu être expérimenté, on constate qu'une réponse positive n'est souvent pas suffisante et qu'un travail de relance doit être fait. Une fois les textes commentés, comme pour la branche « évaluation », il convient d'inviter les auteurs à y réagir.

3.2.2. Estimation du temps de travail nécessaire

Il n'est pas aisé de déterminer précisément ce que la fonction de facilitation nécessite comme temps de travail. Mais il apparaît clairement qu'elle repose sur la place de pivot occupée par le secrétariat de rédaction dans une revue²⁶. On peut affirmer que la rapidité des retours – et donc une évaluation

26 CONTAT Odile et GREMILLET Anne-Solweig, « Publier : à quel prix ? Étude sur la structuration des coûts de

rapide – est liée à un accompagnement appuyé. Bien évidemment, plus de temps est accordé à cette tâche implique son meilleur déroulement – administrer le dispositif en ligne, trouver des contributeurs, démarcher les évaluateurs, favoriser les échanges par exemple. Il est préférable que cette tâche soit effectuée par une personne ayant une certaine connaissance des champs de recherche concernés ou une facilité à comprendre les spécificités d'un champ et/ou une curiosité interdisciplinaire. Cela permet de cerner plus rapidement les évaluateurs et contributeurs potentiels et de comprendre au mieux les cultures disciplinaires spécifiques pour désamorcer d'éventuels problèmes de communications. Dans le cas de l'expérience menée, une activité à temps plein a permis de se consacrer entièrement à dix textes pendant quatre mois, en réduisant les retours d'évaluations à un mois, ce qui est relativement court par rapport à ce qui est généralement pratiqué, et à rechercher activement des contributeurs. Ce temps plein a également été occupé par une recherche périphérique sur les questions d'évaluation. Pour un nombre de textes identique à celui de l'expérience, soit cinq dans chacune des branches, sans se préoccuper de la mise en place du dispositif, on pourrait estimer le temps de travail nécessaire à un minimum de 17 heures par semaine²⁷. La branche « commentaire » ayant demandé davantage de temps, on pourrait en déduire une répartition de 7 heures pour la branche « évaluation » et 10 heures pour la branche « commentaire ».

Recommandations 2

- **L'évaluation ouverte par les pairs et le commentaire ouvert appellent des tâches de secrétariat de rédaction spécifiques.**
- **Un important travail d'explicitation du dispositif doit être mené par le secrétariat de rédaction, les échanges entre auteurs et évaluateurs/contributeurs nécessitent un suivi et doivent être encouragés.**
- **Le démarchage individuel des contributeurs pour la branche « commentaire » est essentiel pour que des interactions se développent effectivement autour des contenus publiés.**
- **Le coût d'un article évalué ouvertement est estimé à 17h soit 448€ ; le coût d'un article ouvert à commentaire est estimé à 24h soit 638€, se situant donc dans une moyenne basse.²⁸**

publication pour les revues françaises en SHS », in *Revue française des sciences de l'information et de la communication*, 7, 2015, (Consulté le 11/02/16), URL : <http://rfsic.revues.org/1716>

27 Dans le cadre de l'expérimentation, on peut considérer que cette tâche de secrétariat dédiée à l'évaluation ouverte a été assurée sur la moitié du temps de travail, soit 17 à 20 heures par semaine.

28 En se basant sur le coût de travail horaire utilisé dans l'enquête citée ci-dessous. Ce coût se situe dans une fourchette basse, un tel dispositif n'implique donc pas de frais particulièrement plus élevés.

3.3. Penser et formaliser la rétribution des contributeurs

Enjeu du crédit des évaluateurs

L'évaluation ouverte par les pairs révèle une question centrale propre à la publication scientifique, au-delà de la certification qualitative des articles publiés. L'évaluation est une part essentielle du processus éditorial, elle est toutefois invisible. L'évaluation ouverte rompt cette invisibilité, elle permet donc de rétribuer les évaluateurs de manière symbolique, sur le même mode qu'une publication. En effet, l'évaluation ouverte permet de notifier la contribution des chercheurs à une évaluation, activité propre à la recherche qui devrait pouvoir être prise en compte dans les activités professionnelles d'un chercheur et mise en valeur. C'est cette problématique qui anime des initiatives telles que Publons²⁹. Ce site internet propose à la communauté scientifique de recenser les évaluations, en collaboration avec les éditeurs. L'intérêt suscité par l'initiative se constate aux plus de 54 000 évaluateurs recensés et aux plus de 16 000 publications faisant évaluer leurs contenus via le site. Comme nous l'a confirmé un correspondant à Publons, l'intuition à la base de la création de ce site était bien de réussir à faire valoir le travail d'évaluation dans les carrières des chercheurs :

« Our founder, Andrew Preston, had plenty of first hand experience with the needs of academics while pursuing his PhD and publishing his own papers. He felt that vision of a hub for post-publication engagement with papers could only succeed if researchers could use it to further their careers. This intuition proved accurate despite asking researchers to use a completely new and unrecognised system being too great a request at first. »

L'existence et le succès de Publons démontrent l'intérêt de pouvoir créditer les évaluateurs de publications scientifiques. De plus, la possibilité d'être crédité pour son travail d'évaluation et de commentaire peut permettre d'attirer un plus grand nombre de chercheurs à ces activités qui sont quelque peu délaissées. Nous avons nous-mêmes constaté la difficulté, confirmée également par la rédactrice adjointe de la revue, de trouver des évaluateurs disponibles. Cela est encore plus problématique pour intéresser des contributeurs. Il reste à savoir comment l'accréditation des évaluateurs/contributeurs peut être intégrée à un dispositif d'évaluation ouverte par les pairs et de commentaire ouvert, sans être externalisé par un service tel que proposé par Publons. De manière générale, comme cela a déjà été mentionné plus haut, c'est un véritable référencement de ces activités et sa prise en compte dans les métadonnées des publications qui apporteraient une solution viable.

29 <https://publons.com>

Le crédit par le remerciement et la mention

Dans le cadre du dispositif expérimenté, différentes solutions ont été mises en place dans l'optique de créditer le travail des évaluateurs et des contributeurs. Dans le cas de la branche « évaluation », les noms et affiliations des évaluateurs sont indiqués sur la page où a eu lieu le processus d'évaluation – le contenu en lui-même n'est dépublié que dans le cas où un texte est rejeté, mais la page ne cesse pas d'exister, la mention aux évaluateurs est donc pérenne. Il est également recommandé à la rédaction de la revue de mentionner les noms et affiliations des évaluateurs dans la version publiée des textes. Cela peut par exemple être intégré comme note de la rédaction. Dans le cas de la branche « commentaire », le même système de mention des contributeurs est adopté sur la page où était publié le texte. Il est de plus suggéré aux auteurs d'ajouter les contributeurs en remerciement dans leur nouvelle soumission d'article – ce qui a effectivement été fait dans les versions corrigées.

La mention en ligne des noms et des évaluateurs et des contributeurs de manière pérenne – sur les pages en question ou sur une page dédiée à cette fin – permet de laisser une trace de cette activité éditoriale et éventuellement pour les chercheurs d'y faire référence si ils le souhaitent. Une fois de plus, les initiatives telles que Publons ou Reviewer Page, laissent à penser que la citation de ces activités sera une pratique de plus en plus courante. A terme, l'idéal serait de pouvoir référencer dans les métadonnées des publications les noms des évaluateurs et contributeurs à un texte, ainsi que des permaliens renvoyant vers le processus d'évaluation assurant la traçabilité de sa première certification scientifique. Cela est envisageable sur des textes finalement publiés, mais la mise en place de telles métadonnées pourrait être envisagée sur les pages où se déroulent les évaluations elles-mêmes pour ne pas négliger le travail d'évaluation ou de commentaire réalisé sur des textes qui ne seraient finalement pas publiés.

Recommandation 3

- **Les évaluateurs et les contributeurs doivent être formellement crédités pour le travail effectué. Cette accréditation doit être pérenne pour pouvoir y faire référence.**
- **Les textes publiés doivent mentionner la manière dont ils ont été évalués et mentionner les noms des évaluateurs, les auteurs dont les textes prépubliés ont été commentés doivent être invités à remercier nommément les contributeurs. La revue doit veiller à ces accréditations.**

- **Toutes les informations liées à l'évaluation devraient être intégrées aux métadonnées des publications et des pré-publications (type d'évaluation, noms, référence électronique).**

OUVERTURES ET PERSPECTIVES

A l'issue de cette expérience, il convient en premier lieu de retenir l'enthousiasme manifesté autour de sa réception. Il existe une importante littérature sur l'évaluation ouverte cherchant à départager ses avantages de ses inconvénients. Si son existence même le suppose, elle ne rend pas compte, à ma connaissance, de l'ampleur de l'intérêt qu'exprime la communauté scientifique sur la question. Outre ce point qu'il est important de rappeler, trois idées générales autour de l'ouverture des processus d'évaluation et de commentaire seront maintenant avancées en guise de conclusion. Puis, le prototype expérimenté sera modélisé pour que d'autres dispositifs semblables puissent s'en inspirer. On peut auparavant signaler ce que la mise en place de cette expérience permet de réaliser de manière générale. Tout d'abord, l'ouverture de l'évaluation et l'ouverture à commentaires sont des moyens de démystifier les pratiques d'évaluation, quelque peu paralysées par les enjeux de carrière qu'elles sous-tendent. Les pratiques d'évaluation classiques se sont en effet imposées par convention et semblent difficile à dépasser historiquement. L'ouverture permet ce dépassement et pourrait permettre de détendre les crispations liées aux enjeux de carrière. D'autre part, il ne faut pas négliger le fait que cette ouverture, qui semble innovante, est en réalité ce que pratique déjà la communauté scientifique dans le face à face, dans le cadre de séminaires, de colloques et de conférences. Il ne faut donc pas s'y tromper : ouvrir les modalités d'évaluation et ouvrir des contenus à commentaires ne fait que rendre visible une partie des activités habituelles et nécessaires de la recherche. L'enjeu est donc important pour tisser des liens à l'intérieur de la communauté scientifique, mais aussi en termes de liens entre la recherche et la société. Dans l'édition scientifique, l'ouverture des modalités d'évaluation permet de rendre visible ce qui existe à la discrétion des éditeurs, or cette visibilité ne peut que simplifier la médiation scientifique. On pourrait ici se risquer à avancer que si cette ouverture met en danger les carrières, c'est qu'un problème existe autour de l'avancement et non autour de la manière dont les chercheurs peuvent et doivent communiquer dans et autour de leurs domaines de recherches.

Les résultats et préconisations formulés à l'issue de cette expérience peuvent permettre de modéliser un dispositif opérationnel d'évaluation ouverte par les pairs et de commentaire ouvert. Selon ce qui a été expérimenté, observé et préconisé dans ce bilan, un tel dispositif s'articule autour de trois

dimensions : un espace éditorial adéquat, la possibilité de commenter/annoter les textes et un accompagnement humain spécifique facilitant le fonctionnement du dispositif. Mais avant de proposer une modélisation du prototype, qui synthétisera ce qui a été présenté jusqu'ici, nous souhaiterions suggérer quelques ouvertures possibles, réflexions issues de la recherche périphérique menée autour de cette expérience. La première, confirmée par des échanges entretenus avec deux chercheurs au sujet de l'expérience, souligne le besoin de mise en place progressive de tels dispositifs. La seconde, induite par cette première remarque, concerne la possibilité d'hybrider les différentes formes d'évaluation pour permettre à la communauté de se familiariser avec ces pratiques. La dernière, relevée par un entretien avec la directrice d'une structure éditoriale en sciences humaines et sociales, souligne l'intérêt de renouveler les modalités d'évaluation également dans le champ de l'édition scientifique de monographies et d'ouvrages collectifs.

Une mise en place progressive

Des échanges informels entretenus au long de l'expérience, à intervalles réguliers, avec deux chercheurs ne participant pas à l'expérience, un point significatif émerge. Il concerne le besoin de mise en place progressive de dispositifs d'évaluation et de commentaire ouverts. Cette remarque, déjà soulevée plus haut, s'articule avec le caractère relativement inédit – en tout cas dans la sphère francophone et les sciences humaines et sociales – des dispositifs d'évaluations ouvertes et de l'ouverture à discussion des contenus éditoriaux en général.

« J'ai déjà eu à faire face à des relecteurs qui ne paraissaient pas du tout qualifiés et ont rejeté des articles pour des raisons qui me paraissaient totalement hors sujet. Je pense donc que la levée de l'anonymat est une bonne chose, surtout pour pouvoir discuter avec les relecteurs. Mais selon moi, le problème est surtout lié à une tradition. Je ne suis pas sûre que tout le monde joue le jeu, même si on voit bien avec ce que vous avez monté que ça fonctionne ! A mon avis, il faut commencer par montrer que ça marche et que ça reste fiable, avant de le généraliser. Pas à pas. » ; « On peut facilement se dire que l'ouverture de l'évaluation, c'est quelque chose qui va se développer. Comme l'*open access* pour les publications scientifiques. Mais il va falloir y habituer ceux qui sont persuadés que rien ne doit changer. »

La question est donc de savoir comment « habituer » avant de « généraliser ». La meilleure solution semble être le « pas à pas » pour prouver la fiabilité du dispositif et continuer à l'améliorer. Comme le relève un des deux chercheurs interrogés, le blocage essentiel relève de la tradition d'une évaluation anonyme qui garantirait une plus grande impartialité. Si l'anonymat permet d'éviter des

réactions agressives comme celle du blocage décrit plus haut – ou en tout cas leur publicité et leur personnification, ce qui est important –, il n'empêche en rien les querelles d'écoles et d'éventuels règlements de compte. On peut même penser qu'il les rend possibles dans la mesure où l'évaluateur est protégé par l'anonymat. Pour passer outre l'effet de tradition, il faut envisager une mise en place progressive des dispositifs d'évaluations ouverts. De plus, il faudrait rendre évident la manière dont le texte a été évalué (ouvert/aveugle), en l'indiquant très clairement, et en favorisant la consultation du processus d'évaluation lorsque celui-ci est ouvert. Les deux points de vue cités semblent bien répondre à la question centrale entourant les débats sur l'évaluation ouverte : il n'y a pas de raison de penser qu'elle est moins efficace et moins garante de qualité que l'évaluation classique. Au terme de l'expérience, nous pensons qu'il importe de favoriser la mise en place de tels dispositifs et laisser la communauté scientifique en observer les effets en situation.

Possibilité d'hybrider les différents modèles

Cette première remarque en amène une seconde : la possibilité d'hybrider les différentes formes d'évaluation et de commentaire ouverts. Dans le cadre de notre expérience, les deux dispositifs ont été mis en place pour des prépublications. Mais on pourrait envisager que des textes déjà publiés soient ouverts à commentaires et que cette ouverture soit accompagnée pour qu'un échange effectif puisse se déployer. Comme on l'a vu sur le carnet de recherche de *VertigO*, lorsque la rédaction ouvre des textes à commentaires, ceux-ci restent peu commentés, il convient donc de permettre à cette dynamique de s'engager en la facilitant. Comme nous avons tenté de le mettre en avant, les enjeux de tels dispositifs sont bien sûr importants pour une meilleure communication scientifique.

D'autre part, pour accompagner une mise en place progressive de tels dispositifs ouverts, on pourrait imaginer que seuls certains textes d'une revue soient soumis à une évaluation ouverte avec les accords des auteurs et des évaluateurs. Le retour sur expérience ne nous indique pas de réticences majeure à ce type d'innovation, à l'exception de l'évaluateur du texte 1 qui a vécu une situation difficile avec l'auteur. On peut donc postuler que de nombreux auteurs et évaluateurs accepteraient d'utiliser un tel dispositif. Cela permettrait une adaptation progressive de la communauté à ce type de pratiques et, encore une fois, lui laisserait l'opportunité d'en constater par elle-même l'efficacité.

Ouvrir aux monographies

Dans cette même dynamique d'introduction progressive des pratiques d'évaluations ouvertes, on peut également envisager une ouverture de l'évaluation des monographies et ouvrages collectifs. Les modalités d'évaluation des périodiques et des ouvrages scientifiques sont très proches, elles diffèrent cependant sur la question de l'anonymat. Dans le processus qui nous a été décrit par la directrice d'une structure éditoriale en sciences humaines et sociales rencontrée dans le cadre de notre recherche périphérique, un comité se prononce sur la pertinence de publication d'un manuscrit avant de le soumettre à expertise. Lorsqu'un manuscrit est accepté, il est ensuite retravaillé par l'auteur en lien avec l'expert qui le relit. Si la première appréciation de pertinence se fait donc en aveugle – puisque l'auteur ne sait pas qui se prononce sur la pertinence ou non du manuscrit proposé – il semble de règle que le travail sur le manuscrit avec un expert se fasse sans la médiation de l'éditeur. Cette ouverture de l'activité de travail sur un manuscrit, qui s'apparente donc bien à de l'évaluation dans le sens où ce sont des corrections qui sont alors demandées, est tout à fait légitime comme l'explique l'éditeur :

« La médiation formelle est facile à faire. Mais ce qui est très difficile à faire, c'est la médiation scientifique. Dans l'exemple que je donne, l'expert sait exactement ce qu'il doit dire à l'auteur : "Pour gagner en lisibilité, je te propose d'écarter tel type d'exemple parce qu'il n'est pas très intéressant, tu pourrais retravailler le lien entre tes deux chapitres, etc..." (...) Tout un ensemble de propositions que nous ne ferons pas par la suite, puisqu'on adoptera plutôt le point de vue du lecteur lambda. Il nous arrive donc très très souvent de mettre la directrice de collection, le lecteur du comité de lecture ou l'expert externe en relation avec l'auteur. Ils ont la même responsabilité et de l'autonomie vis-à-vis de l'auteur. Ils s'engagent vraiment à nous [l'éditeur] aider à rendre un livre lisible.

- Et ces relations se passent bien en général ?

- Très bien. (...) Ces relations se passent très bien à partir du moment où les auteurs sont partants pour retravailler. Si ils ne sont pas partants, ils ne le font pas... »

Ainsi, on pourrait imaginer en prépublication que la révision d'un manuscrit se fasse en ligne, que cet échange puisse laisser une trace publique et que le travail du ou des relecteur.s soit ainsi mis en visibilité. D'autre part, il pourrait être envisagé d'ouvrir les textes finalisés à commentaires. Cela leur offrirait un prolongement post-publication où auteurs et lecteurs pourraient continuer à faire évoluer le texte. Cela existe d'ailleurs déjà, à de hauts degrés d'ouverture, l'exemple typique étant celui du livre de Kathleen Fitzpatrick dont l'ouvrage *Planned Obsolescence*³⁰ a été soumis à une

30 FITZPATRICK Kathleen, *Planned Obsolescence : Publishing, Technology and the Future of Academy*, NYU Press, New York, 2011, 256 p. Voir également: <http://www.plannedobsolescence.net> et <http://mcpress.media-commons.org>

évaluation ouverte par les pairs.

Une synthèse : OPRISM – Open Peer Review facilitation through Social Media

Le prototype expérimental développé par OpenEdition de novembre 2015 à février 2016 peut être modélisé et systématisé en tant que proposition d'outil d'évaluation ouverte par les pairs et de commentaires ouverts. Ce prototype s'appuie sur trois dimensions :

Publication et discussion : une interface de publication de contenus scientifiques à évaluer avec possibilité technique de les commenter,

Annotation : un outil permettant d'annoter les contenus,

Facilitation : un accompagnement humain de facilitation.

Les possibilités d'échange entre auteurs et évaluateurs permises par la dimension ouverte (non-anonyme et en ligne) sur une interface dédiée, facilitées par un travail spécifique de secrétariat de rédaction, donnent son nom au prototype pouvant être systématisé : OPRISM, *Open Peer Review facilitation through Social Media*.

Dans le cadre de l'expérience testant le prototype décrit ici ces trois dimensions sont respectivement le carnet de recherche d'une revue scientifique hébergée sur Hypothèses (<http://hypotheses.org>), l'utilisation du logiciel d'annotation Annotator via Hypothes.is (<http://annotatorjs.org> & <https://hypothes.is>) et un secrétaire de rédaction. Pour décrire cette modélisation, nous verrons d'abord en quoi cette expérience a fait ses preuves dans un environnement représentatif – que l'on peut donc considérer de niveau de maturité technologique TRL6 – et dans quelle mesure un tel dispositif peut être applicable de manière générale sur des plates-formes de publication scientifique dédiées à des articles, des monographies et des ouvrages collectifs. On détaillera ensuite la pertinence des trois dimensions caractérisant le dispositif : le besoin de facilitation, l'interface offerte par Hypothèses et le choix du logiciel développé par Annotator.

OPRISM - prototype expérimenté en milieu significatif

Le dispositif facilité d'évaluation ouverte par les pairs et de commentaires ouverts OPRISM a fait ses preuves dans un environnement représentatif : une publication scientifique à comité de lecture en accès libre, dont les contenus sont systématiquement évalués par les pairs. Le prototype a donc

été déployé en conditions réelles.

Dans ce cadre, deux dispositifs distincts ont été testés, concernant tous deux des prépublications:

- un dispositif d'**évaluation ouverte par les pairs** dans lequel les rapports d'évaluations suivant une grille donnée sont publiés à la suite des textes proposés et des annotations sont publiées dans le corps du texte. Ces rapports et annotations sont signés et un échange peut se développer autour d'eux entre auteurs et évaluateurs.
- un dispositif de **commentaire ouvert** dans lequel des textes non encore soumis à évaluation, sont discutés et annotés de manière non-anonyme en vue de leur amélioration. Un échange peut alors s'engager autour de ces observations.

Les évaluations ouvertes ont relevé différents cas de figures allant de l'acceptation de l'article avec demande de corrections mineures ou majeures au rejet de l'article proposé. Le retour sur expérience collecté auprès des protagonistes ne montre pas que les évaluateurs aient particulièrement nuancé leurs propos en raison de l'ouverture du processus d'évaluation, il montre cependant que de nombreux évaluateurs ont porté une attention particulière à la qualité de leur évaluation pour rendre leurs observations le plus compréhensible possible. Du côté des auteurs, le retour sur expérience montre que l'identification des évaluateurs leur a permis de mieux cerner leurs observations en appréciant leur légitimité et le cas échéant en échangeant avec eux. Bien que l'enjeu soit différent dans les deux parties du dispositif pour l'évolution de l'article – évaluation et commentaire – ces remarques y sont valables de manières égales.

Un modèle généralisable à tous types de publications et un outil à intégrer sur les plateformes de publication

La recherche associée à la mise en place du prototype a montré l'intérêt des éditeurs scientifiques à renouveler les modalités d'évaluation de tous types de publications scientifiques : articles, monographies, ouvrages collectifs. En effet, il semble nécessaire de dynamiser la fonction et les modalités d'évaluation, qui se trouvent par la force de l'habitude, délaissées et marquées par un certain désintérêt. Introduire de l'innovation dans la communication scientifique autour des prépublications d'une part, et permettre aux différents protagonistes – auteurs, évaluateurs, éditeurs – d'apprécier la réflexion des uns et des autres de manière plus collaborative d'autre part s'avère être une possibilité pour dynamiser les pratiques d'évaluations et de commentaires de

prépublications. Il ressort également de cette recherche périphérique que dans le cas des monographies, la mise en relation directe entre évaluateurs et auteurs pour l'amélioration d'un manuscrit est une pratique courante et souhaitable. De fait, l'anonymat qui est la règle dans l'évaluation classique des articles perd tout intérêt dans le cas des monographies quand il s'agit d'entamer une discussion sur la correction d'un manuscrit. De la même manière, l'anonymat du commentaire ouvert ne présente pas d'intérêt particulier. Le processus de commentaire autour d'une pré-publication est pertinent dans la mesure où les uns et les autres peuvent identifier leurs domaines d'expertises respectifs. Enfin, on peut envisager une ouverture à commentaire d'un manuscrit déjà publié sur un espace dédié, où auteurs et lecteurs peuvent échanger des vues sur le contenu publié. Cette possibilité n'a toutefois pas ici été testée.

Dans la mesure où ce prototype est applicable à tous types de publications scientifiques, on peut envisager sa systématisation en tant que module sur des plates-formes électroniques de publication scientifique. Ainsi, éditeurs et rédactions disposeraient d'un outil déjà intégré et opérationnel permettant d'ouvrir les processus d'évaluations, d'ouvrir des contenus à commentaires et enfin de pouvoir conserver et rendre accessibles les différentes versions des publications ainsi que les échanges qui les accompagnent.

Facilitation : la nécessité d'un accompagnement humain

L'expérimentation du prototype met en évidence la nécessaire facilitation du processus. L'outil technique n'est pas suffisant pour que le modèle puisse se déployer de manière efficace, il va nécessairement de pair avec des ressources humaines d'accompagnement, c'est-à-dire une activité adaptée de secrétariat de rédaction. Celle-ci est un travail de facilitation dans la mesure où elle doit s'attacher à :

- **faciliter la compréhension du processus d'ouverture,**
- **faciliter les prises de contacts et les conditions d'échange,**
- **faciliter l'utilisation technique de l'outil.**

Ces activités de facilitation s'imposent dans la mesure où l'ouverture des modalités d'évaluation reste peu commune, elle doit donc être accompagnée d'un point de vue didactique : un travail d'explication du dispositif doit être réalisé. D'autre part, comme dans un dispositif classique d'évaluation, un important travail de relance doit être effectué. Celui-ci est d'autant plus important

dans un dispositif ouvert dans la mesure où il ne convient pas simplement de relancer les retardataires, mais de favoriser l'échange entre protagonistes. Il s'avère donc nécessaire d'inviter personnellement les uns et les autres à interagir une fois que des observations sont publiées. D'un point de vue pratique, c'est également un travail de facilitation technique qui doit être réalisé. Le facilitateur doit lever les obstacles techniques rencontrés par les utilisateurs.

Si le travail de facilitation peut paraître plus important dans le cas de l'évaluation ouverte par les pairs que pour le commentaire ouvert, cela est contredit par l'expérience. Le commentaire relevant davantage de l'activité facultative, sur laquelle aucun engagement formel n'a été pris, le facilitateur doit encourager la communauté scientifique ou le corps social concerné par l'objet de la recherche à intervenir sur le texte. Contrairement à l'évaluation, le commentaire ouvert doit pouvoir lever les barrières de la légitimité scientifique pour s'ouvrir au-delà de la communauté et permettre la potentialité d'un échange social plus large.

Publication et discussion : une interface de publication de contenus scientifiques à évaluer avec possibilité technique de les commenter

Le prototype expérimenté a été mis en place sur un carnet de recherche hébergé sur Hypothèses, dépendant lui-même d'une publication scientifique hébergée sur Revues.org. Partant, on peut considérer que OPRISM pourrait s'appuyer sur Hypothèses, plate-forme de carnets de recherche reconnue et techniquement structurée de manière adéquate :

- **simplicité d'utilisation,**
- **différentes possibilités de structuration en fonction des contenus à évaluer,**
- **possibilité de commenter et discuter.**

La pertinence de l'utilisation d'Hypothèses comme interface d'évaluation et de commentaire ouverts s'articule avec le nécessaire travail de facilitation, en effet la simplicité de publication sur les carnets de recherche d'Hypothèses ne nécessite pas de connaissances techniques approfondies. On peut donc aisément postuler que ce travail est réalisable par quiconque dispose d'une maîtrise moyenne des outils de publication numérique. On peut postuler que cela est le cas de tout secrétaire de rédaction d'une publication électronique. Hypothèses s'appuie sur le système de gestion de contenu WordPress, généralement utilisé comme moteur de blog, distribué sous licence publique générale GNU. D'autre part, la possibilité de structuration d'un site WordPress en

catégories et sous-catégories permet d'envisager, par exemple, d'y reproduire une arborescence chapitrée. Enfin, et cela est le point central, la forme moteur de blog de WordPress permet la publication de commentaires à la suite d'un contenu et permet de publier des réponses à ces commentaires.

Dans le cadre du prototype développé par OpenEdition, le facilitateur prend en charge la publication des premières évaluations, c'est-à-dire de la retranscription d'une grille d'évaluation standard confiée à l'évaluateur – faute de formulaire automatisé. Laisser au facilitateur le soin de retranscrire cette grille dans le cas d'une double évaluation permet de formaliser les retranscriptions de manière cohérente et de les rendre ainsi le plus claires possible pour l'auteur, mais également pour les autres lecteurs. Enfin, le choix d'Hypothèses semble particulièrement pertinent dans la mesure où la plate-forme est d'ores et déjà reconnue et utilisée par la communauté scientifique, en particulier en sciences humaines et sociales : Hypothèses héberge à la fin janvier 2016, 1433 carnets de recherche.

Annotation : un outil permettant d'annoter le contenu ligne à ligne

Pour annoter les prépublications ligne à ligne, le prototype expérimenté par OpenEdition a fait le choix d'utiliser la librairie JavaScript libre Annotator via le plug-in Hypothes.is. Celui-ci fonctionne tel un calque se déployant sur une page internet via le navigateur utilisé, mais on pourrait imaginer une meilleure intégration de l'outil aux plates-formes de publication dans un souci de confort et de rapidité d'utilisation. L'intérêt d'utiliser un tel outil fonctionnant soit sur le navigateur de l'utilisateur soit sur la plate-forme elle-même est de pouvoir afficher les annotations directement où le texte est publié plutôt que de renvoyer vers un document collaboratif de type Framapad ou Googledoc où seraient collectées les annotations. La version d'Annotator développée par Hypothes.is présente d'autres fonctionnalités qui rendent l'outil pertinent dans un modèle d'évaluation et de commentaire ouverts. Il dispose d'une fonction de réponse aux annotations et crée un permalien pour chacune d'entre elles permettant la traçabilité et la pérennité des annotations. Les outils basés sur Annotator permettent également de créer des groupes si besoin est de restreindre le nombre d'utilisateurs. Hypothes.is est un logiciel libre et gratuit. Une alternative viable à l'utilisation d'Annotator pourrait être celle de CommentPress, également libre et gratuit, mais qu'il faudrait obligatoirement intégrer à l'interface.

Un modèle facilité d'évaluation ouverte par les pairs et de commentaire ouvert opérationnel

pour des plates-formes de publication scientifique

Le prototype OPRISM expérimenté par OpenEdition est un modèle viable et opérationnel pour la mise en place d'un dispositif d'évaluation ouverte par les pairs et/ou de commentaire ouvert. Au-delà de ce que l'expérience a directement prouvé sur la pertinence de l'articulation entre les trois dimensions du prototype que sont l'espace adéquat fourni par Hypothèses, les possibilités d'annotations offertes par Annotator et le besoin d'une facilitation humaine, l'expérience démontre la possibilité d'appliquer ce modèle à l'évaluation et le commentaire ouverts de tous types de publications scientifiques. Ainsi, OPRISM pourrait être un outil intégré aux plates-formes de publications scientifiques, permettant aux éditeurs d'ouvrir facilement leurs modalités d'évaluation.

REMERCIEMENTS

Cette expérience conduite par OpenEdition a été financée dans le cadre du projet européen OpenAire2020 et réalisée en partenariat avec le consortium Couperin et la revue *VertigO*. L'auteur tient à remercier les nombreux participants à l'expérience (auteurs, évaluateurs et contributeurs), Eric Duchemin et Jessica Onitsoa Andriamasinoro (revue *VertigO*) pour leur suivi, Emmanuelle Corne (Éditions de la Maison des Sciences de l'Homme) pour l'entretien accordé, tout comme ceux et celles qui lui ont permis de nourrir sa réflexion autour des questions d'évaluation.

Enfin, l'auteur remercie tout particulièrement Pierre Mounier (OpenEdition) pour son suivi, ses réflexions et son engagement autour du dispositif et de la rédaction de cet article.

REPERES BIBLIOGRAPHIQUES

Les références proposées ci-dessous n'ont pour la plupart pas été directement citées dans l'article. Mais elles permettent de contextualiser et d'aller plus loin sur la question de l'évaluation par les pairs et de son ouverture.

AÏT ALI Nawel et ROUCH Jean-Pierre, « Le « je suis débordé » de l'enseignant-chercheur », in *Temporalités*, 18, 2013, <http://temporalites.revues.org/2632>, (consulté le 12/02/2016)

BORDIER Julien, « Evaluation ouverte par les pairs », écrit dans le cadre de cette expérience et disponible sur OpenEdition LAB, <http://lab.hypotheses.org/1453> (consulté le 08/02/2016)

BOURE Robert, « De l'évaluation collégiale à l'évaluation à dominante gestionnaire »,

Communication et organisation, 38, 2010, <http://communicationorganisation.revues.org/1391> (consulté le 11/02/2016)

BROWN Tracey (dir.), *Peer review and the Acceptance of New Scientific Ideas*, 2004, Sense About Science, London, 61 p.

BURHAM C. John, « The Evolution of Editorial Peer Review », in *JAMA*, 263(10), 1990, doi:10.1001/jama.1990.03440100023003

CHARTRON Ghislaine, « Une économie renouvelée de la publication scientifique », in *Perspective documentaire en éducation*, 2006, p.7 <sic_00117798>

CONTAT Odile et GREMILLET Anne-Solweig, « Publier : à quel prix ? Étude sur la structuration des coûts de publication pour les revues françaises en SHS », in *Revue française des sciences de l'information et de la communication*, 7, 2015, <http://rfsic.revues.org/1716> (consulté le 11/02/16)

COULMONT Baptiste, “L’Autolib’ révélatrice de la sociologie postmoderne”, in *Le Monde*, 9/03/2015

DECOURSEY Tom, « The pros and cons of open peer review », in *Nature*, 2006, doi : 10.1038/nature04991

DACOS Marin et MOUNIER Pierre. “Les carnets de recherche en ligne, espace d'une conversation scientifique décentrée”. In *Lieux de savoir, T.2, Gestes et supports du travail savant*, 2010, Albin Michel, Paris, pp.333-352 (disponible sur HAL: <sic_00439849>)

FITZPATRICK Kathleen, *Planned Obsolescence : Publishing, Technology and the Future of Academy*, NYU Press, New York, 2011, 256 p. et <http://www.plannedobsolescence.net>

GODLEE Fiona, « Making Reviewers Visible. Openness, Accountability, and Credit », in *JAMA*, 2002. doi:10.1001/jama.287.21.2762

HARNAD Stevan, « The Invisible Hand of Peer Review », in *Exploit Interactive*, issue 5, 2000. <http://users.ecs.soton.ac.uk/harnad/nature2.html>

KOWALCZUK K. Maria et al., « Retrospective analysis of the quality of reports by author-suggested and non-author-suggested reviewers in journals operating on open or single-blind peer review models », *BMJ Open*, Volume 5, Issue 9, 2015, doi:10.1136/bmjopen-2015-008707

KRONICK David Abraham, « Peer review in 18th-century scientific journalism », in *JAMA*, 263(10), 1990, doi:10.1001/jama.1990.03440100021002

LAMONT Michèle, *How Professors Think, Inside the Curious World of Academic Judgement*, 2009, Harvard University Press, Cambridge/London, 330 p.

MARTIN René, SULS Jerry, « The Air We Breathe : A Critical Look at Practices and Alternatives in the Peer-Review Process », in *Perspectives on Psychological Science*, 4:40, 2009, doi : 10.1111/j.1745-6924.2009.01105.x

NIELSEN Michael, « Three Myths about Scientific Peer Review », in <http://michaelnielsen.org>,

8/01/2009, <http://michaelnielsen.org/blog/three-myths-about-scientific-peer-review> (consulté le 08/10/2015)

PONTILLE David and TORNAY Didier, « The Blind Shall See ! The Question of Anonymity in Journal Peer Review », in *Ada : A Journal of Gender, New Media, and Technology*, No.4., 2014, doi : 10.7264/N3542KVV

SHEMA Hadas, « Introduction to open peer review », juin 2014, <http://blogs.scientificamerican.com/information-culture/an-introduction-to-open-peer-review/>

SHOTTON David, « The Five Stars of Online Journal Articles – a Framework for Article Evaluation », in *D-Lib Magazine*, Vol. 18, numéro 1-2, doi:10.1045/january2012-shotton

SPIER R., « The history of the peer-review process », in *Trends Biotechnol*, 20(8), 2002, doi: 10.1016/S0167-7799(02)01985-6

VAN NORDEN Richard, « Publishers withdraw more than 120 gibberish papers », in *Nature*, 24/02/2014, doi:10.1038/nature.2014.14763

VELTEROP Johannes, « Peer Review – Issues, Limitations and Future Development », *ScienceOpen Research*, 2015, doi : 10.14293/S2199-1006.1.SOR-EDU.AYXIPS.v1

VITALI ROSATI Marcelo, « What is editorialization ? » in *Sens Public*, Janvier 2016, <http://sens-public.org/article1059.html> (consulté le 11/02/16)

WICHERTS Jelte M., « Peer Review Quality and Transparency of the Peer-Review Process in Open Access and Subscription Journals », in *PlosOne*, 29 janvier 2016, doi : 10.1371/journal.pone.0147913

WALSH Elizabeth, ROONEY Maeve, APPLEBY Louis, WILKINSON Greg, « Open peer review: a randomised controlled trial », in *The British Journal of Psychiatry*, 176, 2000, doi: 10.1192/bjp.176.1.47